

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: October 14, 2006, 03:55:44 ; Search time 312 Seconds

(without alignments)
2217.666 Million cell updates/sec

Title: US-10-687-035-1

Perfect score: 3945

Sequence: 1 AAQPARARARTRKLFTHRSSV.....QKLISEDLNMTGHHHHH 748

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2849598 seqs, 925015592 residues

Total number of hits satisfying chosen parameters: 2849598

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: uniprot_sprot:*

2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------|--------------|
| 1 | 3665 | 92.9 | 6995 | 2 | Q96RK2_HUMAN |
| 2 | 3649 | 92.5 | 22152 | 2 | Q8WY77_HUMAN |
| 3 | 3642 | 92.3 | 1148 | 2 | Q9H7S7_HUMAN |
| 4 | 1562.5 | 39.6 | 867 | 2 | Q6ZQW5_HUMAN |
| 5 | 736.5 | 18.7 | 258 | 2 | Q9D1H1_MOUSE |
| 6 | 144.5 | 3.7 | 1408 | 2 | Q27557_METH |
| 7 | 143.5 | 3.6 | 656 | 2 | Q7R3P3_GIALA |
| 8 | 141.5 | 3.6 | 5010 | 2 | Q4MT18_BACCE |
| 9 | 138 | 3.5 | 3443 | 2 | Q8J2M8_MOUSE |
| 10 | 136 | 3.4 | 605 | 2 | Q94K06_ARATH |
| 11 | 135.5 | 3.4 | 1117 | 1 | ENGL1_YEAST |
| 12 | 134.5 | 3.4 | 1526 | 2 | Q6BLB8_DEBNA |
| 13 | 132.5 | 3.4 | 797 | 2 | Q54E23_DICDI |
| 14 | 132.5 | 3.4 | 4981 | 2 | Q2P2L6_MOUSE |
| 15 | 131 | 3.3 | 1230 | 2 | Q54E54_DICDI |
| 16 | 131 | 3.3 | 1447 | 2 | Q54UR7_DICDI |
| 17 | 128 | 3.2 | 879 | 2 | Q4J6P1_SULAC |
| 18 | 128 | 3.2 | 5017 | 2 | Q81FJ0_BACCR |
| 19 | 127.5 | 3.2 | 1011 | 2 | Q2M1P3_HUMAN |
| 20 | 127 | 3.2 | 1166 | 2 | Q4UD04_THEXN |
| 21 | 127 | 3.2 | 1364 | 2 | Q86A18_DICDI |
| 22 | 126.5 | 3.2 | 885 | 2 | Q54V40_DICDI |
| 23 | 126.5 | 3.2 | 2340 | 2 | Q8UZZ7_9POTY |
| 24 | 126 | 3.2 | 1056 | 2 | Q4VWP3_9HERP |
| 25 | 126 | 3.2 | 1818 | 2 | Q54HY5_DICDI |
| 26 | 125.5 | 3.2 | 1011 | 2 | Q54HY5_DICDI |
| 27 | 125.5 | 3.2 | 1011 | 2 | Q54HY5_DICDI |
| 28 | 125 | 3.2 | 1084 | 2 | Q81SK0_BACAN |
| 29 | 125 | 3.2 | 1084 | 2 | Q8EUS5_MYCPE |
| 30 | 125 | 3.2 | 10791 | 2 | Q7U7M8_SYNXP |
| 31 | 124.5 | 3.2 | 1012 | 2 | Q6FQ05_CANCA |

| | | | | | | |
|----|-------|-----|------|---|--------------|--------------------|
| 32 | 124.5 | 3.2 | 2121 | 2 | Q733G8_BACCI | Q733G8 bacillus ce |
| 33 | 124.5 | 3.2 | 2535 | 2 | Q755B8_ASHCO | Q755B8 ashya goss |
| 34 | 124.5 | 3.2 | 5017 | 2 | Q63DP3_BACCI | Q63DP3 bacillus ce |
| 35 | 124 | 3.1 | 856 | 2 | Q6FXG3_CANCA | Q6FXG3 candida gla |
| 36 | 124 | 3.1 | 1442 | 2 | Q967H5_SULTO | Q967H5 sulfolobus |
| 37 | 123.5 | 3.1 | 1736 | 2 | Q3SBB2_EURY | Q3SBB2 uncultured |
| 38 | 123.5 | 3.1 | 1837 | 2 | Q9N5F6_CABRY | Q9N5F6 caenorhabd |
| 39 | 123.5 | 3.1 | 2281 | 2 | Q51R20_GCALI | Q51R20 sapovirus s |
| 40 | 123.5 | 3.1 | 2477 | 1 | PINC_RAT | P04937 rattus norv |
| 41 | 123.5 | 3.1 | 2520 | 2 | Q4MM66_BACCE | Q4MM66 bacillus ce |
| 42 | 123 | 3.1 | 1085 | 2 | Q54DB6_DICDI | Q54DB6 dictyostell |
| 43 | 123 | 3.1 | 1552 | 2 | Q2NH63_9EURY | Q2NH63 methanospa |
| 44 | 123 | 3.1 | 2959 | 2 | Q5UPY0_MIMIV | Q5UPY0 mimivirus |
| 45 | 123 | 3.1 | 3441 | 2 | Q89PB9_BRAJA | Q89PB9 bradyrhizob |

ALIGNMENTS

| RESULT 1 | ID | Q96RK2_HUMAN | PRELIMINARY | PRT | 6995 AA. |
|-----------------------|--|---|-------------|------------|-----------------------------------|
| AC | Q96RK2 | | | | |
| DT | 01-DEC-2001 | | | | Integrated into UniProtKB/TREMBL. |
| DT | 01-MAR-2004 | | | | sequence version 2. |
| DT | 07-FEB-2006 | | | | entry version 13. |
| DE | Mucin 16 (Fragment). | | | | |
| GN | Name=MUC16; | | | | |
| OS | Homo sapiens (Human). | | | | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | | |
| OC | Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; | | | | |
| OC | Homo | | | | |
| OX | NCBI_TaxID=9606; | | | | |
| RN | [1] | | | | |
| RP | NUCLEOTIDE SEQUENCE. | | | | |
| RX | MEDLINE=21347965; PubMed=11369781; DOI=10.1074/jbc.M103554200; | | | | |
| RA | Yin B.W., Lloyd K.O.; | | | | |
| RT | "Molecular cloning of the cal25 ovarian cancer antigen. identification. | | | | |
| RT | as a new mucin, mucin 16." | | | | |
| RL | J. Biol. Chem. 276:27371-27375(2001). | | | | |
| RN | [2] | | | | |
| RP | NUCLEOTIDE SEQUENCE. | | | | |
| RL | Lloyd K.O., Yin B.W.T.; | | | | |
| RU | Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases. | | | | |
| CC | Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms | | | | |
| CC | Distributed under the Creative Commons Attribution-NonDerivs License | | | | |
| CC | EMBL: AF361486; AAK74120.3; -, mRNA. | | | | |
| DR | HSSP; Q9D1H1; 11VZ. | | | | |
| DR | SMR; Q96RK2; 6803-6922. | | | | |
| DR | InterPro; IPR00194; ATPase_a/bcentre. | | | | |
| DR | Pfam; PF01390; SEA; 20. | | | | |
| DR | SMART; SM00200; SEA; 1. | | | | |
| DR | PROSITE; PS00152; ATPASE_ALPHA_BETA; UNKNOWN_1. | | | | |
| DR | PROSITE; PSS0024; SEA; 6. | | | | |
| FT | NON TER | | | | |
| SO | SEQUENCE | | | | |
| | 6995 AA; 744966 MW; 80C797DDBDF33A2B CRC64; | | | | |
| Query Match | 92.9%; | Score | 3665; | DB 2; | Length 6995; |
| Best Local Similarity | 99.9%; | Pred. No. | 8.3e-238; | | |
| Matches | 696; | Conservative | 1; | Mismatches | 0; |
| | | | | Indels | 0; |
| | | | | Gaps | 0; |
| QY | 14 | FTHRSSVSTSTPTGPTVYVIGASKTPASIFGPEASASHLILFTLNTFTITLRYEENMPG | 73 | | |
| DB | 6240 | FTHRSSVSTSTPTGPTVYVIGASKTPASIFGPEASASHLILFTLNTFTITLRYEENMPG | 6299 | | |
| QY | 74 | SRKFNTERVYVQLPLPLRNKTSVGPLYSGCRITLRLPEKDGATGVDAICTHRPPTGP | 133 | | |
| DB | 6300 | SRKFNTERVYVQLPLPLRNKTSVGPLYSGCRITLRLPEKDGATGVDAICTHRPPTGP | 6359 | | |
| QY | 134 | GDREGLVLESLQTLHSITELGPTLDRDRLVNGFTHRSSVPTTSVGVSSEPPFTLNT | 193 | | |

```

DB 6360 GIDRQLVLELSQLTHTSITELPDTLDRSLVYNGFTHRSVPTTSTGVSEEPFTLNFT 6419
QY 194 INNLRVYADMGQPGSLKFNITDNVWKHLSPFORSSLGARTGCRVALALSVKNGAETR 253
DB 6420 INNLRVYADMGQPGSLKFNITDNVWKHLSPFORSSLGARTGCRVALALSVKNGAETR 6479
QY 254 VDLCTCYLQPLSGPGLPIKQVHELSQOHTGITRLGPYSLDKDSLVLNGYNEPDEPPT 313
DB 6480 VDLCTCYLQPLSGPGLPIKQVHELSQOHTGITRLGPYSLDKDSLVLNGYNEPDEPPT 6539
QY 314 TPKPATFTLPLSEATTAMGYHLKTLTNFTISNLQYSPDMKGSATFNSTEGVLQHLR 373
DB 6540 TPKPATFTLPLSEATTAMGYHLKTLTNFTISNLQYSPDMKGSATFNSTEGVLQHLR 6599
QY 374 PLFQSSMGPPFLGCOLISLRPEKGAATGVTTCYHPDPVPGGLDIQOLYWELSQULTH 433
DB 6600 PLFQSSMGPPFLGCOLISLRPEKGAATGVTTCYHPDPVPGGLDIQOLYWELSQULTH 6659
QY 434 GVTOLGPFYLDRLDSLFINGYAPONLSIRGEYOINFIHIVNMNLSNPDPTSEYITLLRDIQ 493
DB 6660 GVTOLGPFYLDRLDSLFINGYAPONLSIRGEYOINFIHIVNMNLSNPDPTSEYITLLRDIQ 6719
QY 494 DKVTTLYKGSQULHDFRFLVNTLTMDSVLYTVKALFSSNLDPSLVEQVFLDKTLNASFH 553
DB 6720 DKVTTLYKGSQULHDFRFLVNTLTMDSVLYTVKALFSSNLDPSLVEQVFLDKTLNASFH 6779
QY 554 WIGSTYQVLDIHYTEMESVYOPTSSSTOHFYLNFTTNLPSQDKAOPGTTNYQRNR 613
DB 6780 WIGSTYQVLDIHYTEMESVYOPTSSSTOHFYLNFTTNLPSQDKAOPGTTNYQRNR 6839
QY 614 NIEDALNOLFRRNSISYFSDCOVSTFRSVPNRHHTGVDSLGNFSLARRVDAIYEEF 673
DB 6840 NIEDALNOLFRRNSISYFSDCOVSTFRSVPNRHHTGVDSLGNFSLARRVDAIYEEF 6899
QY 674 LRMTRNGTOLQNFITLDRSSVLDGYSPPNRNEPLTGNS 710
DB 6900 LRMTRNGTOLQNFITLDRSSVLDGYSPPNRNEPLTGNS 6936

RESULT 2
O8WX17_HUMAN PRELIMINARY; PRT; 22152 AA.
AC O8WX17_
DT 01-MAR-2002, integrated into UniProtKB/TrEMBL.
DT 07-FEB-2006, entry version 16.
DE Ovarian cancer related tumor marker CA125.
GN Name=MUC16;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae;
OC Homo.
OX NCBI_Taxid=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=21646939; PubMed=11786729; DOI=10.1159/000050638;
RA O'Brien T.J., Beard J.B., Underwood L.J., Dennis R.A., Santin A.D.,
RA York L.;
RT "The CA 125 gene: an extracellular superstructure dominated by repeat
RT sequences."
RN Tumour Biol. 22:348-366(2001).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA O'Brien T.J., Underwood L.J., Beard J.B.;
RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonCommercial License
CC EMBL; AF414442; AAL65133.2; -; mRNA.
DR SMR; O8WX17; 21960-22079.
DR Ensembl; ENSG00000181143; Homo sapiens.

```

```

DR HGNC; HGNC:15582; MUC16.
DR InterPro; IPR000194; ATPase_a/bcentre.
DR InterPro; IPR000082; SEA.
DR Pfam; PF01390; SEA; 51.
DR SMART; SM00200; SEA; 1.
DR PROSITE; PS00152; ATPase_ALPHA_BETA; UNKNOWN_1.
DR PROSITE; PS50024; SEA; 11.
SQ SEQUENCE 22152 AA; 2353428 MW; B3E7BDP1997A440 CRC64;

Query Match 92.5%; Score 3649; DB 2; Length 22152;
Best Local Similarity 99.7%; Fred. No. 5.7e-236;
Matches 695; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 14 FTHRSVSTSTPGPTVYLASAKTPASIFGPSASHLLIFLTNFTTNLYEENMPG 73
DB 21397 FTHRSVSTSTPGPTVYLASAKTPASIFGPSASHLLIFLTNFTTNLYEENMPG 21456
QY 74 SRKFVTERVLOGLRPLFKNTSVGPLYSGRLTLRPEKGEATGVDAICTHRDPGCP 133
DB 21457 SRKFVTERVLOGLRPLFKNTSVGPLYSGRLTLRPEKGEATGVDAICTHRDPGCP 21516
QY 134 GIDRQLVLELSQLTHTSITELGPYTLDRDSLVLNGFTHRSVPTTSTGVSEEPFTLNFT 193
DB 21517 GIDRQLVLELSQLTHTSITELGPYTLDRDSLVLNGFTHRSVPTTSTGVSEEPFTLNFT 21576
QY 194 INNLRVYADMGQPGSLKFNITDNVWKHLSPFORSSLGARTGCRVALALSVKNGAETR 253
DB 21577 INNLRVYADMGQPGSLKFNITDNVWKHLSPFORSSLGARTGCRVALALSVKNGAETR 21636
QY 254 VDLCTCYLQPLSGPGLPIKQVHELSQOHTGITRLGPYSLDKDSLVLNGYNEPDEPPT 313
DB 21637 VDLCTCYLQPLSGPGLPIKQVHELSQOHTGITRLGPYSLDKDSLVLNGYNEPDEPPT 21696
QY 314 TPKPATFTLPLSEATTAMGYHLKTLTNFTISNLQYSPDMKGSATFNSTEGVLQHLR 373
DB 21697 TPKPATFTLPLSEATTAMGYHLKTLTNFTISNLQYSPDMKGSATFNSTEGVLQHLR 21756
QY 374 PLFQSSMGPPFLGCOLISLRPEKGAATGVTTCYHPDPVPGGLDIQOLYWELSQULTH 433
DB 21757 PLFQSSMGPPFLGCOLISLRPEKGAATGVTTCYHPDPVPGGLDIQOLYWELSQULTH 21816
QY 434 GVTOLGPFYLDRLDSLFINGYAPONLSIRGEYOINFIHIVNMNLSNPDPTSEYITLLRDIQ 493
DB 21817 GVTOLGPFYLDRLDSLFINGYAPONLSIRGEYOINFIHIVNMNLSNPDPTSEYITLLRDIQ 21876
QY 494 DKVTTLYKGSQULHDFRFLVNTLTMDSVLYTVKALFSSNLDPSLVEQVFLDKTLNASFH 553
DB 21877 DKVTTLYKGSQULHDFRFLVNTLTMDSVLYTVKALFSSNLDPSLVEQVFLDKTLNASFH 21936
QY 554 WIGSTYQVLDIHYTEMESVYOPTSSSTOHFYLNFTTNLPSQDKAOPGTTNYQRNR 613
DB 21937 WIGSTYQVLDIHYTEMESVYOPTSSSTOHFYLNFTTNLPSQDKAOPGTTNYQRNR 21996
QY 614 NIEDALNOLFRRNSISYFSDCOVSTFRSVPNRHHTGVDSLGNFSLARRVDAIYEEF 673
DB 21997 NIEDALNOLFRRNSISYFSDCOVSTFRSVPNRHHTGVDSLGNFSLARRVDAIYEEF 22056
QY 674 LRMTRNGTOLQNFITLDRSSVLDGYSPPNRNEPLTGNS 710
DB 22057 LRMTRNGTOLQNFITLDRSSVLDGYSPPNRNEPLTGNS 22093

RESULT 3
O9H7S7_HUMAN PRELIMINARY; PRT; 1148 AA.
AC O9H7S7_
DT 01-MAR-2001, integrated into UniProtKB/TrEMBL.
DT 01-MAR-2001, sequence version 1.
DT 07-MAR-2006, entry version 19.
DR CDNA FL014503 F18; Clone PLAGE2000132.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae;

```

```

OC Homo.
NCBI_TaxID=9606;
RN [1]
NUCLEOTIDE SEQUENCE.
RC TISSUE=Placenta;
RX PubMed=14702039; DOI=10.1038/ng1285;
RA Oca T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Makamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita R.,
RA Sekine M., Odayashi M., Nishi T., Shishihara T., Tanaka T., Ishii S.,
RA Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahari K.,
RA Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A.,
RA Sudo H., Hosoi T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Kanda K., Yokoi T., Furuya T., Kikawa E., Omura Y.,
RA Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M., Yamazaki M.,
RA Nishimura K., Ishibashi T., Yamashita H., Murakawa K., Fujimori K.,
RA Tanai H., Kimata M., Watanabe M., Hirao S., Chiba Y., Ishida S.,
RA Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T., Kusano J.,
RA Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O., Nomura Y.,
RA Togiya S., Komai F., Hara R., Takeuchi K., Arita M., Imose N.,
RA Mutsaers H., Yuki H., Oshima A., Sasaki N., Aotsuka S.,
RA Yoshikawa Y., Matsunawa H., Ichihara T., Shichata N., Sano S.,
RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
RA Fujimori Y., Komiyama M., Taehiro H., Tanigami A., Fujimura T.,
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Omori Y., Okamoto S.,
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
RA Togaishi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
RT "Complete sequencing and characterization of 21,243 full-length human
RT cDNAs."
RT Nat. Genet. 36:40-45(2004).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
DR EMBL: AK024365; BAB14899.1; -; mRNA.
DR HSSP: Q9D1H1; 11VZ.
DR SMR: Q9H7S7; 956-1075.
DR InterPro: IPR000082; SEA.
DR Pfam: PF01390; SEA; 7.
DR PROSITE: PS50024; SEA; 3.
SO SEQUENCE 1148 AA; 127958 MW; 3861B0D5FDF8A8C CRC64;

Query Match 92.3%; Score 3642; DB 2; Length 1148;
Best Local Similarity 99.6%; Pred. No. 2e-237;
Matches 694; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

OY 314 TPKPATFLPPLSEATTAAGYHLKTLTLNFTISNLQYSPDMKGSATNSREGVLQHLNR 373
DB 693 TPKPATFLPPLSEATTAAGYHLKTLTLNFTISNLQYSPDMKGSATNSREGVLQHLNR 752
OY 374 PLFQKSMGPFYLGQGLISLRPEKQGAATGVDTCTYHDPVPGGLDIQQLYWEISQLTH 433
DB 753 PLFQKSMGPFYLGQGLISLRPEKQGAATGVDTCTYHDPVPGGLDIQQLYWEISQLTH 812
OY 434 GTTQGLFYLDLDSLFINGYAPQNSISGEYQINHIYMNLSNDPSSSEYITLLRDQ 493
DB 813 GTTQGLFYLDLDSLFINGYAPQNSISGEYQINHIYMNLSNDPSSSEYITLLRDQ 872
OY 494 DVTTLTKGSGQADHFRFLVTNLTMDSVLTVAKLFSNLDPSLVEQVFLDKTLNASH 553
DB 873 DVTTLTKGSGQADHFRFLVTNLTMDSVLTVAKLFSNLDPSLVEQVFLDKTLNASH 932
OY 554 WLGSTYQVLVDIHVTEMESSVYQPTSSSTQHFYLPFTITNLPYSQDKAQPGTTVQRNR 613
DB 933 WLGSTYQVLVDIHVTEMESSVYQPTSSSTQHFYLPFTITNLPYSQDKAQPGTTVQRNR 992
OY 614 NIEDALNQLFRNSSIKYFSDQVSTFPGSVPRHHTGVDSLCSNPLARVDRAVLYEER 673
DB 993 NIEDALNQLFRNSSIKYFSDQVSTFPGSVPRHHTGVDSLCSNPLARVDRAVLYEER 1052
OY 674 LMTNNGTQLNFTLDRSSVLVDGYSPNRNEPLTQNS 710
DB 1053 LMTNNGTQLNFTLDRSSVLVDGYSPNRNEPLTQNS 1089

RESULT 4
O6ZQW5 HUMAN PRELIMINARY; PRT; 867 AA.
ID O6ZQW5;
AC O6ZQW5;
DT 05-JUL-2004, integrated into UniProtKB/TrEMBL.
DT 05-JUL-2004, sequence version 1.
DT 21-FEB-2006, entry version 8.
DE CDNA FLJ46845 fig. clone U789300.1946, highly similar to Homo sapiens
DE mcln16 (MUC16).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae;
OC Homo.
NCBI_TaxID=9606;
RN [1]
NUCLEOTIDE SEQUENCE.
RC TISSUE=Uterus;
RA Suzuki O., Sasaki N., Aotsuka S., Shoji T., Ichihara T., Shichata N.,
RA Matsumoto K., Hirano M., Sano S., Nomura R., Yoshikawa Y.,
RA Matsumura Y., Moriya S., Chiba B., Momiyama H., Onogawa S.,
RA Kaeriyama S., Satoh N., Matsunawa H., Takahashi E., Kataoka R.,
RA Kuga N., Kuroda A., Satoh I., Kamata K., Takami S., Terashima Y.,
RA Watanabe M., Sugiyama T., Irie R., Otsuki T., Saito K., Nishikawa T.,
RA Ishii S., Yamamoto J., Isono Y., Kawai-Hiro Y., Saito K., Nishikawa T.,
RA Kimura K., Yamashita H., Matsuo K., Nakamura Y., Sekine M.,
RA Kikuchi H., Kanda K., Wagatsuma M., Murakawa K., Kanehori K.,
RA Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B., Suzuki Y.,
RA Sugano S., Nagahari K., Masuho Y., Nagai K., Isogai T.;
RT "NEO human cDNA sequencing project."
RT Submitted (JUL-2003) to the EMBL/Genbank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
DR EMBL: AK128681; BAC87568.1; -; mRNA.
DR InterPro: IPR000082; SEA.
DR Pfam: PF01390; SEA; 4.
DR PROSITE: PS50024; SEA; 2.
SO SEQUENCE 867 AA; 96201 MW; C7B3033258839622 CRC64;

Query Match 39.6%; Score 1562.5; DB 2; Length 867;
Best Local Similarity 46.3%; Pred. No. 9.9e-97;
Matches 356; Conservative 102; Mismatches 234; Indels 77; Gaps 17;

```

OY 14 FTHRSVSTTPTGPTVYIGASKTPASIFGPSASHLLILFTLNTLTNLRYEENM-WP 72
 DB 101 FTHRFVPTITPTGPTVYIGASKTPASIFGPSASHLLILFTLNTLTNLRYEENMHP 160
 OY 73 GSRKNTTERTVIGLLRPFKNTSVGPYSGCRLTLRLPEKDGATGDAICTHRDPPTG 132
 DB 161 GSRKNTTERTVIGLLRPFKNTSVGPYSGCRLTLRLPEKDGATGDAICTHRDPPTG 220
 OY 133 PGLDREOLVLETSOLTHSTELGPTLDBDSLYVNGFTHRSVPTST----- 180
 DB 221 PGLDREOLVLETSOLTHSTELGPTLDBDSLYVNGFTHRSVPTST----- 180
 OY 181 -----GVYSEEP-----FTLNTLTNLRYEENMHPGSKFNTTNVMMKHLSPFORSS 230
 DB 281 TPASISGPTTASPLVLTINFTLNTLRKEENMMHHPGSKFNTTERTVIGLLRPFKNTS 340
 OY 231 IGARTYGCVALRSVKNCASTRVLLCTYLOPLSGPGLPIQOVFHELSQOHTHTTRIGP 290
 DB 341 VGPLVSGCRLTLRLPEKDGATKVAICTYRPDPKSPGLDRQOLWELISQOLTHSTELGP 400
 OY 291 YGLDSDSLVNGVNEGPDPPTPKPAT-----TFLPPLSE-ATLMAGHLKTLTLNF 343
 DB 401 YTLDRSDSLVNGVNEGPDPPTPKPAT-----TFLPPLSE-ATLMAGHLKTLTLNF 458
 OY 344 TISNLOVSPDM-GKGSATFNSTEGVLQHLRLPFGKSWGPFYLCOLISLPEKDGAT 402
 DB 458 TITNLRKEENMMHHPGSKFNTTERTVIGLLRPFKNTSVGPYSGCRLTLRLPEKDGAT 518
 OY 403 GVDYTCYHPDPVPGGLDIQOLWELISQOLTHSTELGPTLDBDSLYVNGFTHRSVPTST 457
 DB 519 GVDYTCYHPDPVPGGLDIQOLWELISQOLTHSTELGPTLDBDSLYVNGFTHRSVPTST 578
 OY 458 -----LSTRGE-----YQINPHVNNML-SNPDPSTSEYITTLARD 491
 DB 579 STRPGPTVYIGASKTPASIFGPSASHLLILFTLNTLTNLRYEENMMHHPGSKFNTTERTV 638
 OY 492 IODKVTTLKXGSQLDHTFFRCVLYNL--TMDSVLYLVKALFSSNLDPSLV---RQVFLD 545
 DB 639 LOGLRPLFKNTSVGPYSGCRLTLRLPEKDGATGDAICTHRDPPTGGLDRQOLWELISQ 698
 OY 546 -KTLNASFHMLGS-TYQVLDIVHTEMESSVYQPTSSS---STOHPLNFTTNLPLPSODK 600
 DB 699 LSQLTHSTELGPTLDBDSLYVNGFTHRSVPTSTGVYSEEPFTLNTLTNLRYEENMMH 758
 OY 601 APPGTTNVRANKNTEDANOLFRNSISYSFSDCVSTFRFVPRNHHGVUSLGNF-SP 659
 DB 759 GPGSGLKFNITDNVMQHLLSPLFORSLSGARTYGCVALRSVKNCASTRVLLCTYLOP 818
 OY 660 LA-RRUDRVAIYEELFRMTNRGTOLNFTLDRSSVLDGVSPPNRNEPLT 707
 DB 819 LSGPGLPIQOVFHELSQOHTHTTRIGPGLDSDSLVNGFTHRSVPTST 867
 RESULT 5
 OGD1H1_MOUSE PRELIMINARY; PRT; 258 AA.
 AC OGD1H1_MOUSE
 DT 01-JUN-2001, integrated into UniProtKB/TrEMBL.
 DT 01-JUN-2001, sequence version 1.
 DT 07-FEB-2006, entry version 22.
 DE 18-day embryo whole body cDNA, RIKEN full-length enriched library,
 DE clone:111000814 product:hypothetical SEA domain containing protein,
 DE full insert sequence.
 GN Name=111000814Rik;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridae; Muridae; Murinae; Mus.
 OC NCB1_TaxID=10090;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=whole body; STRAIN=C57BL/6J;
 RA MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;

RA Carninci P., Hayashizaki Y.;
 RT "High-efficiency full-length cDNA cloning";
 RM Methods Enzymol. 303:19-44(1999).
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=whole body;
 RX PubMed=16141073; DOI=10.1126/science.1112014;
 RA Carninci P., Kasukawa T., Katayama S., Gough J., Frith M.C., Maeda N.,
 RA Oyama R., Ravasi T., Lenhard B., Wells C., Kodius R., Shimokawa K.,
 RA Davis M.U., Wilson L.G., Aidinis V., Allen J.E., Bailey T.L.,
 RA Ambesi-Impombato A., Anselotti R., Auerbach A., Banno H., Chalk A.M.,
 RA Chiu K.P., Choudhary V., Christoffels A., Clutterbuck D.R.,
 RA Crowe M.P., Dalla E., Dalrymple B.P., de Bono B., Della Gatta G.,
 RA di Bernardo D., Down T., Engstrom P., Fagiolini M., Faulkner G.,
 RA Fletcher C.F., Fukushima T., Furuno M., Futaki S., Gariboldi M.,
 RA Georgakilas N., Gingeras T.R., Gojovic T., Green R.E.,
 RA Guernicchi S., Harbers M., Hayashi Y., Hensch T.K., Hirokawa N.,
 RA Hill D., Hummelbeck L., Iacono M., Ieko K., Iwama A., Ishikawa T.,
 RA Jakt M., Kanapin A., Katoh M., Kawasawa Y., Kelso J., Kikuchi S.K.,
 RA Kitzano H., Kohler G., Krishnan S.P., Kruger A., Kummerfeld S.K.,
 RA Kurochkin I.V., Lareau L.F., Lazarevic D., Lipovich L., Liu J.,
 RA Liuni S., McWilliam S., Madan Babu M., Madera M., Marchionni L.,
 RA Matsuda H., Matsuzawa S., Miki H., Mignone F., Miyake S., Morita K.,
 RA Mottagui-Tabar S., Mulder N., Nakano N., Nakachi H., Ng P.,
 RA Nielsen R., Nishiguchi S., Nishikawa S., Nori F., Ohara O.,
 RA Okazaki Y., Orlando V., Pang K.C., Pavan W.J., Pavoni G., Pesole G.,
 RA Petrovsky N., Piazza S., Reed J., Reid J.F., Ring B.Z., Ringwald M.,
 RA Roost B., Ruan Y., Salzberg S.L., Sandelin A., Schneider C.,
 RA Schombach C., Sekiguchi K., Semple C.A., Seno S., Sessa L., Sheng Y.,
 RA Shibata Y., Shimada H., Shimada K., Silva D., Sinclair B.,
 RA Sperling S., Stupka E., Sugizaki K., Sulana R., Takenaka Y., Taki K.,
 RA Tamura K., Tan S.L., Tang S., Taylor M.S., Tegner J., Teichmann S.A.,
 RA Ueda H.R., van Nimwegen E., Verardo R., Wei C.L., Yang J., Yeh C.,
 RA Yamashita H., Zdobych E., Zhu E.T., Zimmer A., Hilde W., Bult C.,
 RA Grimmond S.M., Teasdale R.D., Liu E.T., Brusic V., Quackenbush J.,
 RA Wahlestedt C., Matlack J.S., Hume D.A., Kai C., Sasaki D., Tomaru Y.,
 RA Fukuda S., Kanamori-Katayama M., Suzuki M., Aoki J., Arakawa T.,
 RA Iida J., Imamura K., Itoh M., Kato T., Kawaji H., Kawagashira N.,
 RA Kawashima T., Kojima M., Kondo S., Kono H., Nakano K., Nimura Y.,
 RA Nishio T., Okada M., Pleasly C., Shibata K., Shirai T., Suzuki S.,
 RA Tagami M., Waki K., Watanabe A., Okamura-Oho Y., Suzuki H., Kawai J.,
 RA Hayashizaki Y.;
 RT "The transcriptional landscape of the mammalian genome";
 RM Science 309:1559-1563(2005).
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=whole body;
 RX PubMed=16141073; DOI=10.1126/science.1112009;
 RA RIKEN Genome Exploration Research Group, and Genome Science Group
 RG (Genome Network Core Team) and the FANTOM Consortium;
 RT "Antisense transcription in the mammalian transcriptome";
 RM Science 309:1564-1566(2005).
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=whole body;
 RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi U., Bono H., Kondo S.,
 RA Nakido I., Oseko N., Saito R., Suzuki H., Yamashita I., Kiyosawa H.,
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojovic T.,
 RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
 RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
 RA Blake J.A., Brad D., Brusic V., Chochia C., Corbani L.E., Cousins S.,
 RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Fretzer K.S.,
 RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
 RA Grimmond S., Guernicchi S., Hirokawa N., Jackson I.J., Jarvis E.D.,
 RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
 RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
 RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
 RA Nagashima T., Naita K., Okido T., Pavan W.J., Petrea G., Pesole G.,
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramchandran S.,
 RA Ravasi T., Reed J.C., Reid D.J., Ring B.Z., Ringwald M.,

RA Sandelli A., Schneider C., Sample C.A., Seiou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verato R., Wagner L., Wahlstedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wysshaw-Boris A., Yamagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavalon M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kikukawa T., Kono H., Nakamura M., Sakazume N., Sato K.,
RA Shitaki T., Waki K., Kawai J., Aizawa K., Arikawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa S.,
RA Miyazaki A., Sasaki K., Sasaki D., Shibata K., Shingagawa A.,
RA Yaunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.,
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs." ;
RL Nature 420:563-573(2002).
[5]
RN NCULEOTIDE SEQUENCE.
RP STRAIN=CS7BL/6J, TISSUE=Whole body;
RC MEDLINE=21065566; PubMed=11217851; DOI=10.1038/35055500;
RX Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arikawa T., Hara A., Fukunishi Y., Kono H., Adachi T., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Caavaent T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochia H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schirral L.M., Staudt F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baladrelli R., Bersh G.,
RA Blake U., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Ralston M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Guettich S., Hill D., Hofmann M., Hume D.A., Kamlya M., Lee N.H.,
RA Lyons P., Marchionni U., Maehira M., Mazzarelli U., Montecoris P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Saeki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Wetz C., Whitteker C., Wilming L.,
RA Wysshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.,
RT "Functional annotation of a full-length mouse cDNA collection." ;
RL Nature 409:685-690(2001).
[6]
RN NCULEOTIDE SEQUENCE.
RP STRAIN=CS7BL/6J, TISSUE=Whole body;
RC MEDLINE=20499374; PubMed=11042159; DOI=10.1101/9r.145100;
RX Carninci P., Shibata Y., Hayatsu M., Sugihara Y., Shibata K., Itoh M.,
RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.,
RT "Normalisation and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes." ;
RL Genome Res. 10:1617-1630(2000).
[7]
RN NCULEOTIDE SEQUENCE.
RP STRAIN=CS7BL/6J, TISSUE=Whole body;
RC MEDLINE=20530913; PubMed=11076861; DOI=10.1101/9r.152600;
RX Shibata K., Itoh M., Aizawa K., Nagasaka S., Sasaki N., Carninci P.,
RA Kono H., Akiyama J., Nishi K., Kusunai T., Tashiro H., Itoh M.,
RA Suni N., Ishii Y., Nakamura S., Hazama M., Nishino T., Harada A.,
RA Yamamoto R., Matsunoto H., Sakaiguchi S., Ikegami T., Kaishiwagi K.,
RA Fujiwara S., Inoue K., Togawa Y., Izawa M., Ohara E., Watabiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsunura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.,
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer." ;
RL Genome Res. 10:1757-1771(2000).
[8]
RN NCULEOTIDE SEQUENCE.
RP STRAIN=CS7BL/6J, TISSUE=Whole body;
RC Arikawa T., Aizawa K., Akinari S., Akimura T., Arai A., Aono H.,
RA Arikawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
RA Hanagaki T., Hara A., Hayatsu N., Hizumoto K., Hirooka T., Hori F.,
RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
RA Kawai J., Kojima Y., Kono H., Kouda M., Koya S., Kurihara C.,
RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ono M.,
RA Okazaki Y., Okido T., Owa C., Saito H., Saio R., Sakai C., Sakai K.,
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shingagawa A., Shitaki T.,
RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,

[illegible]

DR Pfam; PF06739; SBSP; 7.
KM Complete proteome.
SQ SEQUENCE 1408 AA; 152737 MW; 294EC742ABB29F2 CRC64;
Query Match 3.7%; Score 144.5; DB 2; Length 1408;
Best Local Similarity 20.1%; Pred. No. 2.4;
Matches 157; Conservative 99; Mismatches 258; Indels 269; Gaps 40;
QY 58 NRTIINLRKEENMMPSRKRENTT---ERTLOGLRLPFRKTSVGPRLYSCLRLLRLREPD 114
DB 444 NRTVTV-----PDAKKNLGLSDGFLTRPTTISNISVTP----- 480
QY 115 GEATGVDAICTHRPPTGPGDLREBOYLRL-SQLTSHITELGPLYTLDRSLVYNGTHRS 173
DB 481 -----NOTGP-----LTVTLRGNITNCGDSTGWRD---LVING--RT 515
QY 174 SVPTSTGVSEBPTNFTI-----NLRVADMQQ 205
DB 516 AGKWEVSGLETEPEFEFEYTLNNSRAYSVGVNFPFATRVFLGPLIPENLRVTPSGQ 575
QY 206 PGSLKRNITDNMKHLSPFORSSLGARYTGCRIALASVNGAETRY----- 254
DB 576 E-PLKVNVTADLVNVDLDPDSTYAEV--YIDGVLLDSNVTVNASSRTVTSFNRTLAAG 631
QY 255 --DLCTVYLP-----LSGPGGLPIKOVFHELSQOTHG-----ITRL--GPYSLDK 295
DB 632 LVEITINDLEPELVYVMEGRKFIEN--FTLTPGSGAALVTYVSAMITNIDSNPRSYTA 689
QY 296 DSLYLNG-----YNEPGPDEPPTPKPATYFLPLSEATTAMGVH---LKTTL-LNF 343
DB 690 -TIYVGVVDHTKVLNIPGES--TVPFSTSLIPRGVLTISINNNGVGTAVLSEANF 745
QY 344 TISNLOVSPDMGKSGATFSTEGVLOHLRLPLFOKSMQPFYLCQOLSLRPEKQAAAG 403
DB 746 TISNVTVSPEGKSPINVTVAIV-----RNNGDLAG 777
QY 404 VPTTCYHDPVPVPGGLDIQOLYWEISQLTHGVTLQFVYLDRLSLFINGVAPONLSIRGE 463
DB 778 -DPAVTLYLDVA-----WETRTVS-----VPGKSSVLS--FKKELAFPG 816
QY 464 YQINFIH---VNMNLSNPPTSSEYITLLRDIDOKYTTLYKGSQLDHTRFCL-VTN--- 516
DB 817 YRLNLSGTDTVTRVLEPDPTINGF-----NVTPLVGPAPL--SVRASLVNTNPD 865
QY 517 -----LTMSVLYTVKALFSSNIDPSLVEQVFLDKTLNMFHMLG----- 556
DB 866 LVIGFTARLMDGVVVOENIV--SLSPETREIANGTLLTPGNHTVGVNBSKIVRVLR 922
QY 557 -STYOLVDIHVTMESSVYOP-----TSSSSTOHFYLNFT---ITNLPSYSDKAQPGT 605
DB 923 PASITISDLRYT--PSSGSPFLITATATARKNTGVDGNTAVLYINGLAVIDEKVTVGA 980
QY 606 TNYQRKNRIEDALNOLFRRNSI-----KSYFSDCOVSTFRSVP---NRHHT 649
DB 981 -----GRSVQVAFNHTIENAGIYLAGISLTPLDVRVLSRPAISMLSATPLGVSPHRI 1034
QY 650 GVDSLCNFS-----PLARVRVVAIYEFELMTANGTOLQNTIDRS-----SYLVNG 697
DB 1035 IYVALVSTTEBEGSNYTAGLYDGVVYQNTVAVTGPVSIVSFTADISBGRHQTVAIS 1094
QY 698 YSP 700
DB 1095 LSP 1097

RESULT 7
Q7R3F3_GIALA PRELIMINARY; PRT; 656 AA.
AC Q7R3F3;
DT 15-DEC-2003, integrated into UniProtKB/TrEMBL.
DT 15-FEB-2006, sequence version 1.
DT 07-FEB-2006, entry version 16.
DE Q7R3F3_GIALA_79919_77949.

OS Giardia lamblia ATCC 50803.
OC Eukaryota; Diplomonadida; Hexamitidae; Giardia.
OX NCBI_TaxID=164922;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=WB C6;
RA Morrison H.G., McArthur A.G., Adam R.D., Aley S.B., Gillin F.D.,
RT Olsen G.J., Sogin M.L.;
RT "Draft sequence of the Giardia lamblia genome."
Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC -----
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NonCommercial license
CC -----
CC EMBL; AACB0100015; EAA41889.1; -; Genomic DNA.
DR HSP; P10081; 1PDK.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0016787; F:hydrolyase activity; IEA.
DR GO; GO:003676; F:nucleic acid binding; IEA.
DR GO; GO:000166; F:nucleotide binding; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR01545; DEAD/DEAH N.
DR InterPro; IPR01650; Helicase_C.
DR Pfam; PF00270; DEAD; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELIC; 1.
KW ATP-binding; Helicase; Hydrolyase.
SQ SEQUENCE 656 AA; 73379 MW; EEF73D219C01978B CRC64;
Query Match 3.6%; Score 143.5; DB 2; Length 656;
Best Local Similarity 21.1%; Pred. No. 0.89;
Matches 114; Conservative 63; Mismatches 185; Indels 179; Gaps 24;
QY 271 IKQV-----FHELSQTHGIRTLGPRYSLDKSLVYN---GVNEGPDEPPTPKPATYF 321
DB 116 IREVPRIATFEDLSR-----PPDDPE-VQNTYRAKYFOPLPIOKALP----- 160
QY 322 LPLSEATTAMGVHAKTL-----TLNFTISNLO-----YSPDWG----- 355
DB 161 -----TGMVGYDILLACQTSQSGTKCAFIILIRIATEKJLKTMSGHHEDRFRN 213
QY 356 -KGSATFNSTEGVLOHLRLPLFOKSMQPFYLCQOLSLRPEKQAAAGVDTTCYHDP 414
DB 214 AKGSRAYPFC--IIMSPTRRELVOQTAKASWML-----SYGTSILTRVAYGGP 259
QY 415 VPGGLDIQOLYWEISQLTHG-----VTOLGFVYLDL--DSLFINGVAPONLSIR 461
DB 260 SSPQRDALQMGCDILVATPGRLLDPIKQGVETTVTRFVPEFCRMLDMGEFQIRDL 319
QY 462 GEYQINFIHVNMLSNPPTSSEYITLLRDIDOKYTTLYKGSQLDHTRFCLVTNLTM 520
DB 320 HELPRHHSVQ--DPSNPDLTHQ-----IERQTLFSAATFRE-----IKULAMGF 363
QY 521 -----SVLTVYKALFSSNIDPSLVEQVFLDKTLNMFHMLGSTYOLVDIHVTMESSVYQ 575
DB 364 LKQDRIVSITVQIGSSN--PMLAQRVVLYVERSNKRLRL----- 401
QY 576 PTSSSTOHFYLNFTITNLPSYSDKAQPGTMYQRKNRIEDALNOLFRRNSIKSYFSDC 635
DB 402 -TEYITGNADANNLIEV-YGTDKEQ-----DPSLSTVSSTAISSEFENRA 446
QY 636 QVSTFRSVNRRHHTGDSL-----CNFSPLARV-----DRVAIYEBFL 674
DB 447 KNAITY---PDSIAARHDTLANKDGIIOYQIVTFNFRSEADRIFRYEDMKRYVAVIHGM 503
QY 675 RMTKNGTOLQNTLDRSSVLYVDGYSPNR-----NEPLTGNASADIQSGRSSIE 723
DB 504 TOKERENMLKYFKAGRTNLTIGTVAQRGLDIPVRLVLYNDLPGVVDYTRIGRTGA 563

OY 724 G 724
DB 564 G 564

RESULT 8
OAMT18_BACCE PRELIMINARY; PRT; 5010 AA.
ID OAMT18_BACCE
AC 04MT18; 02-AUG-2005, integrated into UniProtKB/TrEMBL.
DT 02-AUG-2005, sequence version 1.
DT 07-FEB-2006, entry version 3.
DE Cell surface protein.
GN ORFNames=BCE_G9241.1620;
OS Bacillus cereus G9241.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus;
OC Bacillus; Firmicutes; Bacillales; Bacillaceae; Bacillus;
OX NCBI_Taxid=269801;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=G9241;
RX PubMed=15155910; DOI=10.1073/pnas.0402414101;
RA Hoffmaster A.R., Ravel J., Rasko D.A., Chapman G.D., Chure M.D.,
RA Marston C.K., De B.K., Sacchi C.T., Fitzgerald C., Mayer L.W.,
RA Maiden M.C.J., Priest F.G., Barker M., Jiang L., Cer R.Z.,
RA Riltone J., Peterson S.N., Weyant R.S., Galloway D.R., Read T.D.,
RA Popovic T., Fraser C.M.;
RT "Identification of anthrax toxin genes in a Bacillus cereus associated
RT with an illness resembling inhalation anthrax."
RL Proc. Natl. Acad. Sci. U.S.A. 101:8449-8454(2004).
CC -! CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC -----
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
CC EMBL; AAEK01000008; EAL15315.1; -; Genomic DNA.
CC GO; GO:0005727; C:extrachromosomal circular DNA; IEA.
CC DR InterPro; IPR000850; Adenylate_kin.
CC DR InterPro; IPR000194; ATPase_a/Dcentre.
CC DR InterPro; IPR001434; DUF11.
CC DR Pfam; PF01345; DUF11; 26.
CC DR TIGRFAMs; TIGR01451; B_ant_repeat; 35.
CC DR PROSITE; PS00113; ADENYLATE_KINASE; UNKNOWN_1.
CC DR PROSITE; PS00152; ATPASE_ALPHA_BETA; UNKNOWN_1.
CC KW Repeat.
CC SQ SEQUENCE 5010 AA; 521891 MW; B00914651BP832DD CRC64;

Query Match 3.6%; Score 141.5; DB 2; Length 5010;
Best Local Similarity 18.5%; Pred. No. 26;
Matches 153; Conservative 123; Mismatches 328; Indels 223; Gaps 36;

OY 10 RRLKTHSSVSTSPGPTVYLGASKTPASIFGSSASHLLIFLTANFTTNRYEN 69
DB 3658 QALITTKTSNPTVDIGTILYISEVKIIGV-----DANIIIFT-----D 3699
OY 70 MWPGSRKFTTERVYQGLLRP--LFKNTSVGLPVSGLRLLTTL--RPEKDGATG- 119
DB 3700 SIPAGTTFVPSDVTINGVQPDTPNENGISIGTIPNSKTLIFQVQTNPPTEFEIVVQ 3759
OY 120 VDAICTHRDPPTGPGGLDR-----EQLYELSLQ--LTHSITEL 154
DB 3760 SSAMVQYVSIPTAPVNRSATSNIVTSLQNNANISIKQADVTAFSIGQNIITYTTLQNI 3819
OY 155 GGYTLD-----RDSLIVNGFTIRSSVP--TSTGVSESEPTLNTINN 196
DB 3820 GTVPANNTLFINIPEGTIFIEDSISNNVIOGAPNENGITLGTIQPEVTYISFOVL 3879
OY 197 LRYMADMGQPSLKFNITDNVKKHL---SPLFORSSLGARYT---GCRVIALNSVKNG 249
DB 3880 TSI-----PPGNTVINISDTSYEYQIEPSPSIIQRSLSNATVTEVRTANVSAALSA-NR 3933

OY 250 AETRYDLCTYQLQPLSGP-LPIKOVFHELSQLTHGTRLGPSLDKSLVINGYNEPGR 308
DB 3934 SITRIGQIITTVAVTANAGVPIITWTL--LIDAIAGTFTIPNSILVDGI-----PRP 3984
OY 309 DEPTTPKPAATFPLPPLSEATATMAGYHLKTLTL-----NETISNLOYSPD--MGKGA 359
DB 3985 NENPIGTATLNIILP--NNTIIVFQVNVVSIPONNININAVIHVEYQPPSPPISE 4041
OY 360 TPNSTEGVQLHRLRPLFKSSMGFPYLGQ-----LISLRPEMDGAATGDTCTYHP 412
DB 4042 TTSSNTWNIQFIDALLITKSKANTLIANIDETIEYTVI---QNNGSIT--TNSIFPT 4094
OY 413 DPVGRGLDIQQLYWELSQTHGVQLGFRVYLDRLDSLFIN-----GYAPQNLSTR 461
DB 4095 DTIEDGA-----VPIPGSVIYNNVTLPEADNRIGISIRINIAS 4132
OY 462 GEYQINFHIVMNLN--NPDETSERY-----TLARDIODKVTTLVYKGSQALHD---TF 509
DB 4133 QATITTFQVSVTNLPAVNPPTNTANIYVDFIENPDPAFIQKSTSTNTTFVQINDADIVSL 4192
OY 510 RFLVTLNLTMDSLVTVKALFSS-NLPSLVEQVFLDKTLNASFHWLGSSTYQLVDIHYTE 568
DB 4193 KTVDLTSVITGIDILYTTTLANTGNTDATAV--VFTDNI PG-----GTFP----IDGSV 4240
OY 569 MESSVYQPTSSSGTGHF-----LNFITTNLYSQDKAQPGTTVQRKKNIE 616
DB 4241 LVNNIPQLANPSTGILVGTATPANSIPVTSVYVIALPASGHVONOSTSRYTIN----- 4295
OY 617 DALNQLFNSSSIKSYFSCQVSTFERSVNRHHTGVDSLNCNPSPLARVDRAVYEEFLRM 676
DB 4296 -----GEOQISTSNITFEVTITAVTAKTRPI-QYADLQIITTYTISI 4338
OY 677 TRNGT-QLQNFLLD-----RSSVLVDGYSFNRNEPLTNSAD 712
DB 4339 TNNGNIGVENIIVTDIIPANTSFIENSIVYVNGNAPNDNPISGIDID 4385

RESULT 9
Q8JZM8_MOUSE
ID Q8JZM8_MOUSE PRELIMINARY; PRT; 3443 AA.
AC Q8JZM8;
DT 01-OCT-2002, integrated into UniProtKB/TrEMBL.
DT 01-OCT-2002, sequence version 1.
DT 07-FEB-2006, entry version 19.
DE MUC4.
DE Name=Muc4;
GN Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=129/SvJ;
RX MEDLINE=22079426; PubMed=12084055;
RA Desseyn J.-L., Clavierau I., Laine A.;
RT "Cloning, chromosomal localization and characterization of the murine
RT mucin gene orthologous to human MUC4."
RL Eur. J. Biochem. 269:3150-3159(2002).
CC -----
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
CC EMBL; AF441786; AAM66254.1; -; mRNA.
CC DR EMBL; AF520423; AAM66746.1; JOINED; Genomic DNA.
CC DR EMBL; ENSMUSG00000035656; Mus musculus.
CC DR MGI; MGI:2153525; Muc4.
CC GO; GO:0007160; P:cell-matrix adhesion; RCA.
CC DR InterPro; IPR005533; AMOP.
CC DR InterPro; IPR006210; EGF_3.
CC DR InterPro; IPR00742; EGF_3.

```

DR InterPro: IPR006209; EGF_like.
DR InterPro: IPR013032; EGF_like_reg.
DR InterPro: IPR003886; NIDO.
DR InterPro: IPR001846; VWF_D.
DR Pfam: PF03782; AMOP; 1.
DR Pfam: PF00008; EGF; 1.
DR Pfam: PF06119; NIDO; 1.
DR Pfam: PF00094; VMD; 1.
DR SMART: SM00723; AMOP; 1.
DR SMART: SM00181; EGF; 1.
DR SMART: SM00539; NIDO; 1.
DR SMART: SM00216; VMD; 1.
DR PROSITE: PS00856; AMOP; 1.
DR PROSITE: PS00022; EGF_1; UNKNOWN_1.
DR PROSITE: PS00026; EGF_3; 2.
SQ SEQUENCE 3443 AA; 365216 MW; 88CC32D3226F632B CRC64;

Query Match 3.5%; Score 138; DB 2; Length 3443;
Best Local Similarity 20.8%; Pred. No. 25;
Matches 182; Conservative 92; Mismatches 324; Indels 276; Gaps 40;

QY 16 HRSSTSTSTGTPV-----YLQAS-----KTPASIF-----GPSASHLLIFLINF 59
DB 1391 HTGSGTSSNPQSTTTPVTTSTLSASSRDQVQTESSQRTIPPGETTTSHASSLSSGP 1450
QY 60 TITNL--RYENMMPGSRKFTTERVLOQ-----LIRPLFKRTSV--GPLVSGRLT 107
DB 1451 STTNMLTSTSTQITSGDTRHTTAIVTQSTPATVQTSILTSQMSVSAPIITSSQLS 1510
QY 108 LIRPERK-----DGEATGDAICTHRPPTGPGDLREQLYELSQL----- 147
DB 1511 TLRGQHTGSKGTSNHOQTITTPVVTSTPSAT-----SRDQIQTESSLRITSPGETTTS 1566
QY 148 -THSTELGPTTLDDSLVNGFTIRSSVPTTS-----TGVSSE--EPFLNFTIN-NL 197
DB 1567 HASSLSSSGPST-----TNMLTSTSTQITSGDTRHTTAIVTQSTPATVQTSILTS 1619
QY 198 RYMADMGPGSLKFNITDNVMGHLSPFORSLGARYGC----- 238
DB 1620 RNMSIVSRP-----ITST--HKSLTLPQROHTSKGTSNPQTTTPKMTTSPSATS 1670
QY 239 -----RVIA-----LRSYKNGAETRYDL----- 256
DB 1671 HDLIERETSSQRTISGETTTSYAPIMSSASPTTHMLSTSTSTSTVDTRHTTLMTQ 1730
QY 257 -----LCYLIQP-----LSGPGLPIKQVFEHS---QQT-----GITKLG 289
DB 1731 GSTPATVQVSPSSKNMSTVSTPTST-HKSLTLPQSGHTGSKGTSNPQTTTPVTTST 1789
QY 290 PYSLDKDSLVLNGVNEP--GPDEPTTPKPATTFPLPSEATTANGYHLKTLTNFTISN 347
DB 1790 PSATTRDQIQTESSQRTISPEETTTSHAPSMISLAP---STTN-----LSTTSS 1838
QY 348 LQYSPDMKGSA--TFNSTEGVLQ-----HLIRPLPQKSSMGPFLYG 387
DB 1839 OSTSDGHTHTTAIVRQSGSTPATVQVSLSSQNTSTVSTMTSTHKLSTLPQSGHTGSM--- 1895
QY 388 COLISLRPEKGAAGVDTTCYHPDPVPGGLDQQLYWEISQLTHGYTQLGFFVYLDSDS 447
DB 1896 -----GTSNPNQTTT-----P-----EVTSTSPSAT--SYDQIQET 1926
QY 448 LPIINGVAPONTLSIREYOINFIIVMNLNPNPT-----SEEVYTLRLDQDKYTTLYK 501
DB 1927 SFQRTISP-----DETTTS--HAPSMNSAPSTHKLSTASTETITSV-DTRHTTAITTE 1978
QY 502 GSQLDTRFCLVNTLMDSV-LVTYKALFSSNLDPSLVEQVFLDKTNLASFHMLGSTYQ 560
DB 1979 GSTTLANT-----QSLTSPSSQMSVSAPIITSSQLSTLRQSGHTGSKGTSNHOQTITTP 2033
QY 561 LVDIVTEMESSVYPTSSSTQOHFYLANFTTNLFLYQDKAOPGTTNFORKNRIED--- 617
DB 2034 VVTSTSPATSRDQIQTESSLRITISPDGTTSHASSMSSSPNTTHLITTSSTESTSV 2093

```

```

QY 618 -----ALNOLFRNSSIKSYFSDCOVST-----FRSPVNRHHTGVDLSNFS 658
DB 2094 DTGSHVITTHGSLTATVQSLFRSSQMSMTVSMPTSSQELTSLPQROHTG--SMETSS 2151
QY 659 PLARRVDRVAIYEFELMKTRNGTLOLQNFYLDRSS 692
DB 2152 QPONITPVTVTSTLTSFRSGSTELQTMWGTSS 2185

RESULT 10
ID 094K06 ARATH PRELIMINARY; PRT; 605 AA.
AC 094K06;
DT 01-DEC-2001, integrated into UniProtKB/TrEMBL.
DT 01-DEC-2001, sequence version 1.
DT 07-FEB-2006, entry version 15.
DE tRNA adenylyltransferase-like protein.
GN Name=t22J18.17; Ordered locusName=Atlg22660;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OK NCBI_TaxID=3702;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Lam B., Southwick A., Karlin-Neumann G., Nguyen M., Miranda M.,
RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamuya A., Kawai J.,
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA Sakurai T., Satou M., Seki M., Shim P., Yamada K., Shinozaki K.,
RA Ecker J., Theologis A., Davis R.W.,
RA Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Nguyen M., Karlin-Neumann G., Southwick A., Tripp M., Miranda M.,
RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamuya A., Kawai J.,
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA Sakurai T., Satou M., Seki M., Shim P., Yamada K., Shinozaki K.,
RA Ecker J., Theologis A., Davis R.W.,
RA Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
CC -! SIMILARITY: Belongs to the tRNA nucleotidyltransferase/poly(A)
CC polymerase family.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivative License
CC -----
CC EMBL: AF370489; AAK43866.1; -; mRNA.
CC DR EMBL: BT000361; AAN15680.1; -; mRNA.
CC DR TAIR: Atlg22660; -.
CC DR GO: GO:0005739; C:mitochondrion; IDA.
CC DR InterPro: IPR012277; Poly(A).
CC DR InterPro: IPR002646; PolyA_pol_reg.
CC DR PANTHER: PTHR13734; PolyA_pol; 1.
CC DR Pfam: PF01743; PolyA_pol; 1.
CC DR Nucleotidyltransferase; RNA-binding; Transferase.
CC KW SEQUENCE 605 AA; 68954 MW; 2DB2778E9368A18 CRC64;

Query Match 3.4%; Score 136; DB 2; Length 605;
Best Local Similarity 20.6%; Pred. No. 2.5;
Matches 99; Conservative 67; Mismatches 185; Indels 130; Gaps 20;

QY 113 KDGEATGDAICTHRPPTGPGDLREQLYELSQLTHSTITELGPTTLDDSLVNGFTNR 172
DB 152 RDEEVQG-DVIERNP-----QSGHLEATL-----RIYQWIDFVNLRSSEYEN 197
QY 173 SSVPTSTGVVSEPFITLFTINLRVYADMGQPGSL-KFNITDNVMGHLSPFORSS- 230
DB 198 SRIFPMKGTAKDAFRDLITNLFLVINGAVEDLTERGIDDKSGKIYTPAPKATF 257
QY 231 -----LGARYGCVIALRSYKNGAETRVLLCTYLOPLSPGGLPIKQVFEHL 278
DB 258 LDDPLRVLAVERFGARFGFTLDEELKEAASSSEVRVAL-----GKISRERIGNEI 308

```

QY 279 SQQTHG-----ITRLGPVSLDKSLYINGVNERGPDEP-----PTTKP 317
 Db 309 DLMISGNGPVAIVTLSDIKLFSVVPALPSSAEPSPENCGLSOSYLEAMSLIKTPRP 368
 QY 318 -----ATTFLP-----PLSEATTAMGYHLK 337
 Db 369 GKFSGEQRALAYAMFLPFRKTVYKDKTKGSIPIVNHIFKPSMRKTSDAETVNHQIT 428
 QY 338 TLTNFTISNLQYSPDMGKSGATFNSREGVLQH-----LLRLPLFKSSMGPPYLGCOLIS 392
 Db 429 TERFSLIPSLSEVKKQVDELDTW--AADILEHMKSTILNDEVIPATSKIRVLG--FL 483
 QY 393 LRPEPD-----GAATGVDTTCYHPDPVPGGLDIQQLYHLSQL--THGVTQGLFYVLD 446
 Db 484 LRDIQDFMRVSLTLTLISLATV--DGSNDHODIGQLDLQLEMRRTYLTVEATIHGLD 541
 QY 447 SLF-----INGVAPONLS-IRG-----EYQIFHIVMNLISNPPTSSEVITLLRDIQD 494
 Db 542 KIMDAKPLVNGREIMQIAELKKGSRILIREWQK--LITWQLAYPMTAECEKEMKRDIXA 599
 QY 495 K 495
 Db 600 K 600

RESULT 11
 ENGI YEAST STANDARD; PRT; 1117 AA.
 AC P53753;
 DT 01-OCT-1996, integrated into UniProtKB/Swiss-Prot.
 DT 07-MAR-2006, entry version 35.
 DE Endo-1,3(4)-beta-glucanase 1 precursor (BC 3.2.1.6) [Endo-1,4-beta-glucanase 1] (Endo-1,3-beta-glucanase 1) (Laminarinase-1).
 GN Name=DSE4; Synonyms=ENGI; Ordered locus names=YNR067C; ORFNames=N3547;
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxID=4932;
 RN 11
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RX MEDLINE=97313269; PubMed=9169873;
 RA Philippsen P., Kleine K., Boehlmann R., Duesterhoeft A., Hamberg K., Hegemann J.H., Obermaier B., Utrera-Tarazu L.A., Aert R., Albertmann K., Altmann R., Andre B., Baladron V., Ballesta J.P.G., Becam A.-M., Beinhauer J.D., Boskovic J., Buित्रago M.J., Buserreau F., Coester F., Crouzet M., D'Angelo M., Dal Pero F., De Antoni A., del Rey F., Daignon F., Domdey H., Dubois E., Fiedler T.A., Fleig U., Floeth M., Fritz C., Galliardin C., Garcia-Cantalejo J.M., Glandsdorf N., Goffeau A., Guelener U., Herbert C.J., Heumann K., Heuss-Neitzel D., Hilbert H., Hini K., Itagui Housaïni I., Jaquet M., Jimenez A., Joniaux J.-L., Karpfinger-Hartl L., Lanfranchi G., Lapingle A., Levesque H., Lyck R., Maftah M., Mallet L., Maurer C.T.C., Messenguy F., Mewes H.-W., Moestl D., Naer F., Nicaud J.-M., Niesenhal R.K., Pandolfo D., Pierard A., Pivrandi E., Plantan R.J., Pohl T.M., Punnelle B., Rebischung C., Remacha M.A., Revelle J.L., Rime M., Saitz J.E., Sartorello F., Scherens B., Sen-Gupta M., Soler-Nira A., Urbanus J.H.M., Valle G., Van Dyck L., Verhasselt P., Viendeele F., Vissers S., Voet M., Volckaert G., Wach A., Wandut R., Wedler H., Zollner A., Hant J.;
 RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome XIV and its evolutionary implications.";
 RL Nature 387:93-98(1997).
 RN 12
 RP FUNCTION, SUBCELLULAR LOCATION, AND GLYCOSYLATION.
 RX PubMed=12455695;
 RA Baladron V., Ufano S., Duenas E., Martin-Cuadrado A.B., del Rey F., Vazquez de Aldana C.R.;
 RT "Eng1p, an endo-1,3-beta-glucanase localized at the daughter side of the septum, is involved in cell separation in Saccharomyces cerevisiae.";
 RL Eukaryot. Cell 1:774-786(2002).

RN [3]
 RP LEVEL OF PROTEIN EXPRESSION.
 RX MEDLINE=22923965; PubMed=14562106; DOI=10.1036/nature02046;
 RA Ghemmaghami S., Huh W.-K., Bower K., Howson R.W., Belle A., Dephoure N., O'Shea E.K., Weissman J.S.;
 RT "Global analysis of protein expression in yeast.";
 RL Nature 425:737-741(2003).
 CC -1- FUNCTION: Involved in the dissolution of the mother-daughter septum during cell separation.
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,3- or 1,4-linkages in beta-D-glucans when the glucose residue whose reducing group is involved in the linkage to be hydrolyzed is itself substituted at C-3.
 CC -1- SUBCELLULAR LOCATION: Cell wall. Localizes asymmetrically to the daughter side of the septum.
 CC -1- PTM: Glycosylated.
 CC -1- MISCELLANEOUS: Present with 64 molecules/cell.
 CC -1- SIMILARITY: Belongs to the glycosyl hydrolase 81 family.
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NonDerivs license
 CC -----
 CC EMBL: Z71682; CA96349.1; -; Genomic_DNA.
 CC PIR: S63399; S63399.
 CC Germonline: 143412; -;
 CC Ensembl: YNR067C; Saccharomyces cerevisiae.
 CC GenomeReviews: Y13139_GR; YNR067C.
 CC BioCyc: SCER-528-01:SCER-528-01-005268-MONOMER; -;
 CC LinkHub: P53753; -;
 CC GO: GO:0030428; C:cell septum; IDA.
 CC GO: GO:0009277; C:cell wall (sensu Fungi); IDA.
 CC GO: GO:0007109; P:cytokinesis, completion of separation; IEPI.
 CC InterPro: IPR005200; Glyco_hydro_81.
 CC Pfam: PF06139; Glyco_hydro_81; 1.
 CC Complete proteome; Glycoprotein; Glycosidase; Hydrolase; Signal.
 FT SIGNAL 1 16
 FT CHAIN 17 1117
 FT FT Endo-1,3(4)-beta-glucanase 1.
 FT FT /FTID=PRO_0000012133.
 FT FT COMPBIAS 342 345
 FT FT COMPBIAS 365 370
 FT FT COMPBIAS 376 383
 FT FT Poly-Ser.
 FT FT CARBOHYD 138 138
 FT FT N-linked (GLCNAc...) (Potential)
 FT FT CARBOHYD 186 186
 FT FT N-linked (GLCNAc...) (Potential)
 FT FT CARBOHYD 223 223
 FT FT N-linked (GLCNAc...) (Potential)
 FT FT CARBOHYD 259 259
 FT FT N-linked (GLCNAc...) (Potential)
 FT FT CARBOHYD 280 280
 FT FT N-linked (GLCNAc...) (Potential)
 FT FT CARBOHYD 303 303
 FT FT N-linked (GLCNAc...) (Potential)
 FT FT CARBOHYD 307 307
 FT FT N-linked (GLCNAc...) (Potential)
 FT FT CARBOHYD 393 393
 FT FT N-linked (GLCNAc...) (Potential)
 FT FT CARBOHYD 533 533
 FT FT N-linked (GLCNAc...) (Potential)
 FT FT CARBOHYD 886 886
 FT FT N-linked (GLCNAc...) (Potential)
 SQ SEQUENCE 1117 AA; 121064 MW; 87F13A07E42B0AD1 CRC64;
 Query Match 3.4%; Score 135.5; DB 1; Length 1117;
 Best local Similarity 21.1%; Pred. No. 6.9;
 Matches 156; Conservative 94; Mismatches 255; Indels 233; Gaps 39;
 QY 4 PARBARFKLTHRSSVSTSTPGPYVLA-----SKTPAS--IFGPAASHL 51
 Db 206 PYSYSGEKIIRP---SLTSNKTITITISRTMAAATGDSFIASIPASTLFPENSTOD 263
 QY 52 LILFTNFTTNLYEEN-----MWPGRKNTTERV-----LOGLRLP 91
 Db 264 LVQ-TLASTTASPAVPSNRQTILSPVSIVSTSPISPIVSNITENGSSPSSLSTVSPV 322
 QY 92 FKNTSVGLVGCCTLLRPEKGEARTGVDAICHRPPTPGGLDRELYLELS-QLTHS 150
 Db 323 YPSSSTGNIL---LSSLFTVDSST-----PVSGTLD--TIYVSSSQATIS 365
 QY 151 ITBGLPYLDRDSLTVNGFTTRSSVPTTSTGVSEEPFTLLFTTNLRYMADMGQPGSLX 210
 Db 366 SSSSSRQTKISS-----SLSTSTSTATTENSTT--TIVNLFNAVSTDEPPIV- 414

CC

```
OY 211 NNIDNNKHHLLSPLEFQRSLGARATGCAVIALRKYNGNAERVDLLCTYIQLPDSRGLE 270
Db 415 FDRSEPNM-----SIADGVSN-----DGP-IO 435
OY 271 IKQVHEH-----SQQTHGTRRLGPSLDK-----DSLINGYNEPGRDEPP 312
Db 436 TNKFYTNLIYVGSQSPAF--VYPSLMKTTSSSGYFAVQHTTYDQSYGSGDSSGNAEYL 493
OY 313 TTPKRAITFLPLSEATTMAGHKLKTLTLNFTISNLQYEPDCKGSGATFRSTEGYL---- 368
Db 494 VNP-----LGIYAVVFSASNDSMTMQVDENMTLSTRTVVISBN 533
OY 369 ---OHLLRPLFOKSSMGPFYLCQQLSLRPEKKGATGVDTTCTYHPDPGPGLDIQOLY 425
Db 534 DSSNLTLEPLVG--GMG-FATGIYHSLN-AKIGSGVGRNTIVSESSNLQOI----LK 585
OY 426 WELSGUITHGVTLQGFVLDRLSLFNGVAPQNLISRGVQINFH-----YNNNLSPND 479
Db 586 YRIT-LLNGVTWL-CYVIGPDDLSTNDP---SLVSESEYEIKASAGVDGLIQLAVA-PS 639
OY 480 PRSSF-----YIT--LALDIQKVTLYK-----GSQL-----HDTFR 510
Db 640 ETDYEVFYDQAAGMYTNPKLGVSQSGSTAYTEFYSYTTGSESASGSTMIPALPHSSFS 699
OY 511 FCLVNLNLTWDSVLVYTKALFFSSNLDSLVEQVFLDTYNASFMHLSSTVO----- 560
Db 700 DIMQDYITQIQLASTTKGYMNGVLTLSLQ----FSTSLRQIWSLFRMSQLGSLNLEYK 755
OY 561 ----LVDIHVTMESSVYQPTSSSSSTQHFYINFTITNLPY-----SQDKAOPGTTN 607
Db 756 EQLQLAEVANSBELQVSISSISGLNT--YILGKVIDKYSYILLTVSEIIODEASTKT- 812
OY 608 YQRNRKRIEDALNQLFRN 625
Db 813 ----LENIKSAFDILLQN 826

RESULT 12
O6BL88 DEBHA
AC O6BL88; PRELIMINARY; PRT; 1526 AA.
DT 16-AUG-2004, integrated into UniProtKB/TrEMBL.
DT 16-AUG-2004, sequence version 1.
DT 21-FEB-2006, entry version 19.
DE Debaryomyces hanseni1 chromosome F of strain CBS767 of Debaryomyces
DE hanseni1.
DE OrderedLocustNames=DEBHA0F167429; (Torulaopora hanseni1).
OS Debaryomyces hanseni1 (Yeast)
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Debaryomyces.
CX NCBI_TaxID=4959; [1]
RN
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=ATCC 36239 / CBS 767;
RX PubMed=15229592; DOI=10.1038/nature02579;
RA Dujon B., Sherman D., Fischer G., Durens P., Casaregola S.,
RA Lafontaine I., de Montigny J., Marck C., Neugeglise C., Talla E.,
RA Goffard N., Frangoul L., Aigle M., Anchaud V., Babour A., Barbe V.,
RA Barnay S., Blanchon S., Beckerich J.-M., Beyne E., Bleyaesten C.,
RA Bolariame A., Boyer J., Catlicchio L., Confantolieri F., de Daruvar A.,
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Giropoli A.,
RA Hantreva F., Henniquin C., Jauniaux N., Joyet P., Kachouri R.,
RA Kerrest A., Koszul R., Lemaître M., Lésur I., Ma L., Müller H.,
RA Nicoud J.-M., Nikolek M., Oztas S., Ozier-Kalogeropoulos O.,
RA Pellenz S., Potier S., Richard G.-F., Straub M.-U., Suleau A.,
RA Swennen D., Tekala F., Wesolowski-Louvel M., Weeshof E., Wirth B.,
RA Zenoun-Meyer M., Zivanovic Y., Bolotin-Fukuhara M., Thierry A.,
RA Bouchier C., Caudeon B., Scarpeilli C., Gallardin C., Weissenbach J.,
RA Winkler P., Souciet J.-L.;
RT "Genome evolution in yeasts.";
RL Nature 430:35-44(2004).
CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
```

| | | |
|---------------------------|--|-----|
| CC | Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms | |
| CC | Distributed under the Creative Commons Attribution-NonDerivs License | |
| CC | ----- | |
| DR | EMBL; CR382138; CAG89403.1 - ; Genomic_DNA. | |
| DR | GO; GO:0005254; F:ATP binding; IEA. | |
| DR | GO; GO:0005488; F:binding; IEA. | |
| DR | GO; GO:0004672; F:protein kinase activity; IEA. | |
| DR | GO; GO:0006488; P:protein amino acid phosphorylation; IEA. | |
| DR | InterPro; IPR011989; ARK-like. | |
| DR | InterPro; IPR000357; HEAT. | |
| DR | InterPro; IPR000719; Prot_kinase. | |
| DR | InterPro; IPR008271; Ser_thr_pkin_AS. | |
| DR | InterPro; IPR002290; Ser_thr_kinase. | |
| DR | InterPro; IPR001680; WD40. | |
| DR | Pfam; PF002985; HEAT_2. | |
| DR | Pfam; PF00069; Kinase_1. | |
| DR | Pfam; PF00400; WD40_3. | |
| DR | PRINTS; PR00320; GPROTEINRPT. | |
| DR | ProDom; PD000001; Prot_kinase; 1. | |
| DR | SMART; SM00320; WD40_1. | |
| DR | PROSITE; PS50077; HEAT_REPEAT; 1. | |
| DR | PROSITE; PS50011; PROTEIN_KINASE_DOM; 1. | |
| DR | PROSITE; PS00108; PROTEIN_KINSE_ST; UNKNOWN_1. | |
| DR | PROSITE; PS50082; WD_REPEATS_2; 1. | |
| DR | PROSITE; PS50294; WD_REPEATS_REGION; 1. | |
| KW | ATP-binding; Complete proteome; Kinase; Nucleotide-binding; Repeat; | |
| KW | Serine/threonine-protein kinase; Transferrase; WD repeat. | |
| SQ | SEQUENCE 1526 AA; 174291 MW; DF5CB86F03820D9C CRC64; | |
| Query Match | 3.4%; Score 134.5; DB 2; Length 1526; | |
| Best Local Similarity | 19.4%; Pred. No. 13; | |
| Matches 159; Conservative | 95; Mismatches 285; Indels 279; Gaps 34 | |
| QY | 115 GEATGVDAICTHRPDTPGGLDREQLYLELSQTHTSIELGP----- | 156 |
| DB | 297 GDPSVDITLNDYDKCFPGFFETLYPMSELNNNDLFVIPSNDNLTPSDLKEIKTY | 356 |
| QY | 157 -YTLDRSLVYNGFTHRSVSPTSTGVVSEEPFLN-----FTINNLRKM | 200 |
| DB | 357 SYDKIADALGFNVSNDRSSLSQNSKFV---PLMLNLKMPKYNTVKPTVTENN--YL | 411 |
| QY | 201 ADMQPGRL-KENTIDNMKKHLISLPQ----- | 227 |
| DB | 412 ---QQGLLIILNLFSLMKFPKOLSKIKACELIALSERVNDCKLDRCIPYLCLIID | 467 |
| QY | 228 -----RSSIGARVYT--CRVIATRSGVGASTRVLDLCITYLOPLSGPQL | 269 |
| DB | 468 EYMSSSTINYQNPNQNTLSNFSSSVACALTSI-----TLLMCSYINPIN---- | 518 |
| QY | 270 PIKOVFHELSOOTHGITRLGPSYLDKDSLVLNYGNVEPGDEPPTPKPATYFLPLSEAT | 329 |
| DB | 519 -VLMFSEYLLPKLHALISIPREDKNLIKLT-----LAACLPLYANVS | 561 |
| QY | 330 TAMGHKLTL-----TLNFIISNIQSPDMKGSAATENS-----TEGYLQHLL | 372 |
| DB | 562 KKFMMMSSTIFKNDVLDKDLNRSLSKXPLDDKENVSDSYNSPIRKQOLDSDENLASKL | 621 |
| QY | 373 ---RPLFGKSSMGPPYLGCOLISLRPKDGATGVDT-----CTYHPRP----- | 414 |
| DB | 622 TDVNEMVAKISLVNNIMPLCOFF-----GVDKTNDIILLPHLITYLDSNYELRIA | 670 |
| QY | 415 -----VGPGLDIQQLYMWELSQLTHGYTOLG-----FYULDRDSLPIINGYA | 454 |
| DB | 671 FLSSITLGIQPFVGV--LSFEQYILPLLQTLGDLEQFVIAKLVELFYCVVRRL-INPKS | 727 |
| QY | 455 PON-LISIRGEYOINFHVNNMLNPDPSSSEYITLLRD-IODKYTTLYK-GSQLDHTF | 509 |
| DB | 728 EFNALSIYEKL-----TSSIKITLLPMEWRQSVCILIALSIDNLLDAD | 772 |
| QY | 510 RFLVLTNLTMSVLTYYAKALSSNLDPSLVEV-----FLDKTLNLSFHHLGSTRY | 559 |
| DB | 773 KYCFYLPYIVKGLVYDVNTIMNNTLYPSITKYPFLSKOIYNLAITWSNASNKSLFPOEKS | 832 |


```

QY 560 QLV-----IHTEMESSVQPTSSSTQHVFNTINLPYQDKAQPNTNQR 610
DB 833 SVFNLPKAPTKLIPYSTNKGKSVYIPKSKNG--PSLNNTNSNIPUSPEDKO-----WVL 884
QY 611 NKRNIEDALNOLFRNSSIKSYSPDCOVSTFRSVPRHHHTGVSL-----CNFSPLARRVD 665
DB 885 KLSKSGVLDNRDLMLKPILRDIYHSSKS-----NTSPKADDFELPKDINTP--RVVF 936
QY 666 RVAIIEEPLRMTRNGTQLONFLDRSSVLVDGYSPPNRNEPLTGSNADIQHSGGRSSLEG 725
DB 937 FEVCYKSEPFSSGSGTAETNF-----ESV-----HTLSNKKDEDSITGLNSLILP 982
QY 726 REEQLIS-----BEDLNMHTGHHH 747
DB 983 NFKVKVSLQTVQANVFGEFLTSTHDSSEFNASSSHHHH 1020

RESULT 13
ID 054E23_DICDI PRELIMINARY; PRT; 797 AA.
AC 054E23;
DT 24-MAY-2005, integrated into UniProtKB/TrEMBL.
DT 24-MAY-2005, sequence version 1.
DE 21-FEB-2006, entry version 7.
DE Hypothetical protein AAC1.
GN Name=AAC1; ORFNames=DD80201568;
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AX4;
RA Eichinger L., Pachebat J.A., Gloeckner G., Rajandream M.-A.,
RA Tuccang R., Bertram M., Song J., Olsen R., Szatranski K., Xu Q.,
RA Tunngal B., Kummerfeld S., Madera M., Konfortov B.A., Rivero F.,
RA Banier A.T., Lehmann R., Hamlin N., Davies R., Gaudet P., Fey P.,
RA Pilcher K., Chen G., Saunders D., Sodergren E., Davis P.,
RA Kerthouar A., Nie X., Hall N., Anjard C., Hemphill L., Bacon N.,
RA Farbrother P., Desany B., Just E., Morio T., Roat R., Chucher C.,
RA Cooper J., Haydock S., van Driessche N., Cronin A., Goodhead I.,
RA Muzny D., Moutier T., Pain A., Harper D., Linday R.,
RA Hauser H., James K., Quiles M., Mohan M.B., Saito T., Buchrieser C.,
RA Wardoper A., Felder M., Thangavelu M., Johnson D., Knights A.,
RA Louisege H., Mungall K., Oliver K., Price C., Quail M.A.,
RA Urushihara H., Hernandez J., Rabbinowitz E., Steffen D., Sanders M.,
RA Ma J., Kohara Y., Sharp S., Simmonds M., Spiegler S., Tivey A.,
RA Sugano S., White B., Walker D., Woodward J., Winckler T., Tanaka Y.,
RA Shalaby G., Schleicher M., Weinstock G., Rosenthal A., Cox B.C.,
RA Chisholm R.L., Gibbs R., Loomis W.F., Platzer M., Kay R.R.,
RA Williams J., Dear P.H., Noegel A.A., Barrell B., Kuspa A.;
RT "The genome of the social amoeba Dictyostelium discoideum.";
RL Nature 0:0-0(2005).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs license
CC -----
CC EMBL; AAF10100264; EAL61460.1; -; genomic_DNA.
CC DR EMBL; AAF10100264; EAL61460.1; LRR.
CC DR InterPro; IPR001611; LRR.
CC DR InterPro; IPR007091; LRR_RNinh.
CC DR InterPro; IPR003590; LRR_RNinh_sub.
CC DR Pfam; PF00560; LRR_1; 2.
CC DR PRINTS; PR00019; LRR_RNinh.
CC DR Hypothetical protein; Leucine-rich repeat; Repeat.
CC KW SEQUENCE 797 AA; 89497 MW; 0CCB0676C5C581FF CRC64;
SQ

```

```

QY 6 RARRTKLTFRSSSVSTTSTPGTPVYLGAASKTASIPGSAASHLLILFTLNFTTNLR 65
DB 3 RTSNRNEATAKLSSSTTTTITTNKYNNMAAES-----SKSTIISLIL 48
QY 66 YEENMPGSRKFNTERVYGLRPLFRKNTSVGLYSGLRLLRPEKGEA-----TGV 120
DB 49 PKYSLEPNNDNDT-----NSSTRP--NKQOKLKSNESTSTTTTTTPI 92
QY 121 DAICHRDPFGPGGLDR---EQYLIELSQLTHSTIELGPTLUDSDSLVNGFTTHSSVPT 177
DB 93 TTTTITTTTTTTTTPLSKYNNVLYIEKQEKONL---PTTETETTTITPTLTITTTTTT 148
QY 178 TSTGVSEB-----PFTL-----NFTINN 196
DB 149 TTTTITTTQIQTSTTSTINHPRLIQKEITFLVELGSLNARKVKCKMKVNGCVEN 208
QY 197 LR-YWADMGQPSLKFNITNDVMKGLSPLFORSSLGARYTGCRVIALRSYKNGAETRV 255
DB 209 LNIYFTDILASVYK-HYSEVFKVSLNSDYFHLQSV-----SFLNGAKNSIS 254
QY 256 LLCTYLOPLSGP-----GLPIKQVHELSQQTGHTGLRGLPSLDDKS 297
DB 255 YSEFENNVILPFIENVRYNQTLENFTIKGFPITRINKSSQOL-----LPSTGLYHST 309
QY 298 LYINGVNERGPDDEPTPKP-----ATFLPLPSEATTAGYHLKT----- 338
DB 310 SVSPSPPPPPPPPPQIQPTTPTAAPTSTTAIVASTTTTAAGQTLNNNNNNNNIPKGLNY 369
QY 339 -LTINFTISNQSPDMKGSAT---FNSTEGVLOHLRLPLFKSMGPPYLACQLIS- 392
DB 370 YLTNNFKLKNLTK-NIGLDSRDKPFPSSLSGNNVLTETLIICDNGD--EGQQLLSV 426
QY 393 -----LPEKDGAAATGVDTTCTTHPDVPGGLDIQQLYWEISQLTHGVTLQGF 440
DB 427 ILIKMLKVLRLLELQKN---OFTNSAYVILNKV---LSCQQLQLETLNLSN----- 473
QY 441 YVLDRLSLFI--NGAPQNTLSIRGEYQINFIHVMNLSNPPTSEVITLLRDIDKYTT 498
DB 474 -RIEQLGIMKDGFG-RNKSLE---EFLPSKRLGNTDSV-----DFGKSITS 517
QY 499 LKYSQQLHDT-----FRFCLVTNLTMDSVLYTVKALFSSNLDPSLVEQVPLDKTL 548
DB 518 L-----DLHDSVSGSKQSLKIGLSOYLKFNESITSLNLSNMHIGSNALISLSPVNGTL 573
QY 549 ---NASFPLMGSTYQL-----VDIHTEMESSVYQPTS-----SSSTQHF-- 585
DB 574 KFDVLSFYKINSNFGDLHLVSSLINHSIHISLSQSNQIDWNTSAITLSQLFNSSRLPSP 633
QY 586 ---YLN-----FTITNLPSYQDKAQPNTNQRKRNIEDLNLQLF 623
DB 634 FKYNLSGNKIGIGLKKLINDLSKYSKTHIITNNNDNNNNKKNENKSKIK--NLSF 690
QY 624 RNSS-----IKSYSPDCQVSTFRSVPRHHHTGVDSLGNFSPLARRVDVAIYEEFLMNT 677
DB 691 NNNNNNPKYIKTKYHND--TPIITIKTINSNGKL-----LEBEVQV-----INS 734
QY 678 NNGTQLONFLDRSSVLVDGYSPPNRNEPLTGSNADIQHSGGRSSLEG-----PRFQK 730
DB 735 SINDTN-QINDNI-----NNENNLTETISLDL---SSNSPLEVSKVIGLIPRYYSK 778
QY 731 L 731
DB 779 L 779

RESULT 14
Q2PZL6_MOUSE PRELIMINARY; PRT; 4981 AA.
ID Q2PZL6_MOUSE
AC Q2PZL6;
DT 24-JAN-2006, integrated into UniProtKB/TrEMBL.
DT 24-JAN-2006, sequence version 1.
DT 07-MAR-2006, entry version 3.
DE Fat4.

```

```

GN Name=Fat4;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN [1]
RP NCBIROTIDE SEQUENCE.
RA Qi C., Zhu Y.T., Hu L., Zhang Z., Rao S.M., Zhu Y.-J.;
RT "Identification of Fat4 as the candidate tumor suppressor gene in
RT breast cancers through random chromosome deletion.";
RL Submitted (NOV-2005) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Membrane; single-pass type I membrane
CC protein (By similarity).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonCommercial License
CC -----
CC EMBL: DQ286572; AB888946.1; -; mRNA.
CC Calcium; Cell adhesion; EGF-like domain; Membrane; Repeat;
CC Transmembrane.
KW SEQUENCE 4381 AA; 540343 MW; 5AC7C40F5C25A500 CRC64;
SQ
Query Match 3.4%; Score 132.5; DB 2; Length 4981;
Best Local Similarity 20.1%; Pred. No. 1e+02;
Matches 165; Conservative 107; Mismatches 299; Indels 249; Gaps 40;
OY 40 ASITGPAASHLLIFLTINFTTNKRVENMPPGSRKNTTER-VIOGLRLPLKQNTSV 97
DB 1587 ASALVPSQ-----LTVNLIVSATDQPR-----RRKSTTELYVILQGLDGPVFTQKY 1634
OY 98 GPLVSGCGLTLRLPEKDEATGVDAICTHRDPPTGPG----- 134
DB 1635 -----ITLL-----KKEGPICGTNVISIEASPGSGEPVEYIVSVACEKTVGRLEPTI 1683
OY 135 -----LDREQ-----LYLEISQUTHSITEIGPYTLDRDLSLYNGFTHRSSVPTTS 179
DB 1684 GRTGVIGTAAILDEGQACLYL-----VDVVAIKSSAPR--TORAVEITLL 1730
OY 180 TGVVSEEPF-----TLNFTNNLRVAMDGQGLKFNITDVMNKLSPLRQBSLGRY 235
DB 1731 QDINDNPVPFPTDITDLTVEE-----NIGDG-----SKINQLTAMDMD- 1768
OY 236 TGCRIYALSVNGAET--RVDLCTCTYLOPLSGPCLPKQYFHELSQOHTGTRLGPSYL 293
DB 1769 EGAAALVYALISGADBSFRID-----PESGDLIAIKRLDRERKSKYSLVRAAD-GL 1820
OY 294 DKDSLVLNGYNPEGPDEPPTTPKPAATFLPPLSEATTAGYHLKTLT-----INFTISN 347
DB 1821 QSDMRINITISDVNDHTPRPSRPVYSF--DIPEDTTGSLVAAILATDDSGVNGEISY 1878
OY 348 LQYSPDKMGKATFNSTGCVLOHLRLPLFOKSSMGPFLGCOLISLREKGAATGVTT 407
DB 1879 V-VERDDDDGVFFLLVTVGF-NLTRALDYETQQ--YY-----ILTVRAEDDG--GQSTT 1927
OY 408 -----CTYHPD-----PYGPGDLOQLYWEISQUTHGV-TQLGYYV 442
DB 1928 IRAYNIIDVANDNPVFSMSYSTSLMNLPLGSTV---LVFNVTDDADGVNGLSISYI 1983
OY 443 LDRDSLFLNGVAPQWLSIRGEYOINFIYVNMNLSNPDPSSBYITLLRIDP-----K 495
DB 1984 ASGDSL-----GQPAVDKGVKLKTLKALDRBSQGFYVLVIOVHDLPPPTSR 2030
OY 496 VTTLYKGS-----QLHDTFRFLCYVTLT-----MDSLVLYVKALPSSNLDPSLVQYFL 544
DB 2031 FTSTAQVSIILLDNDVNDPMPFLSPKLTLYI PRNTPTIDTVVFRKAQADPDGSPNSYIEYTL 2090
OY 545 DK-----TLNAGFWHLG-----STYOLVDIHVEMSSVYOPTSSSTQ--HFYL 587
DB 2091 NPSGKRSIGITIDGVALTGELDRREVSNTSL-----TVVATDKQKQPLSSSTEVVAVL 2145
OY 588 NFTITLPEYDQ-----DKAQPGTTVYQNRKNIEALNQLFENSSIKSYFSDQCVST 639

```

```

DB 2146 DINDNNEVFAQMYRVOIKENILTGTDIIQVSAADNDEGTNGOVRIGVCG----- 2196
OY 640 FRSPVNRHHT-GVDSLGNFSPILARRVR--VAIYEELMTNRTGTOQLNFTLDRSSVLVD 696
DB 2197 -----NTHQFRIDSVYGALTAVAKSIDRETTPTVTLVQATIDRSGSPRTSCYVAITLLD 2251
OY 697 -----GYSPNNEPPLTNGNSADIQHSGGRSLIEGP 725
DB 2252 MNDFVPVPELSPYSVNPENIGTLPRAILQVARDDDQGP 2291
RESULT 15
ID 054B54 D1CDI PRELIMINARY; PRT; 1230 AA.
AC 054B54;
DT 24-MAY-2005, integrated into UniProtKB/TrEMBL.
DT 07-FEB-2006, entry version 1.
DE Hypothetical protein.
DE ORENames=DD80184069;
GN Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostellium.
OX NCBI_TaxID=44689;
RN [1]
RP NCBIROTIDE SEQUENCE.
RC STRAIN=AX4;
RX PubMed=15875012; DOI=10.1038/nature03481;
RA Eichinger L., Pachbat J.A., Gloeckner G., Rajandream M.A., Xu O.,
RA Suedinger R., Berriman M., Song J., Olsen R., Safranek K., Xu O.,
RA Tungal B., Kummerfeld S., Madera M., Kontorov B.A., Rivero F.,
RA Bankier A.T., Lehmann R., Hamlin N., Davies R., Gaudet P., Fey P.,
RA Pilcher K., Chen G., Saunders D., Sodergren E.J., Davis P.,
RA Kerhoun A., Nie X., Hall N., Anjard C., Hemphill L., Bacon N.,
RA Farbrother P., Desany B., Just E., Morio T., Rose R., Churcher C.M.,
RA Cooper J., Haydock S., van Drieseche N., Cronin A., Goodhead I.,
RA Muny D.M., Moutier T., Pain A., Lu M., Harper D., Lindsay R.,
RA Hauser H., James K.D., Quiles M., Madan Babu M., Salto T.,
RA Buchrieser C., Wardrop A., Felder M., Thangavelu M., Johnson D.,
RA Knights A., Louised H., Mungall K.L., Oliver K., Price C.,
RA Quail M.A., Urushihara H., Hernandez J., Rabinowitch B., Steffen D.,
RA Sanders M., Ma J., Kohara Y., Sharp S., Simmonds M.N., Spiegler S.,
RA Tivey A., Sugano S., White B., Walker D., Woodward J.R., Winckler T.,
RA Tanaka Y., Shaulsky G., Schleicher M., Weinstein G.M., Rosenthal A.,
RA Cox B.C., Chisholm R.L., Gibbs R.A., Loomis W.F., Platzer M.,
RA Kay R.R., Williams J.G., Dear P.H., Noegel A.A., Bartell B.G.,
RA Kuapa A.;
RT "The genome of the social amoeba Dictyostelium discoideum.";
RL Nature 435:43-57(2005).
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonCommercial License
CC -----
CC EMBL: AAF01000263; EAL61520.1; -; Genomic DNA.
CC InterPro: IPR002110; ANK.
DR InterPro: IPR002017; Spectrin.
DR Pfam: PF00023; Ank; 3.
DR Pfam: PF00904; Involucrin; 2.
DR PRINTS: PRO1415; ANKYRIN.
DR SMART: SMO0248; ANK; 2.
DR PROSITE: PS50297; ANK_REPEAT; 1.
DR PROSITE: PS50088; ANK_REPEAT; 1.
KW ANK repeat; Hypothetical protein; Repeat.
SQ SEQUENCE 1230 AA; 139239 MW; 6F9B094941E6BD2C CRC64;
Query Match 3.3%; Score 131; DB 2; Length 1230;
Best Local Similarity 18.5%; Pred. No. 16;
Matches 127; Conservative 96; Mismatches 244; Indels 218; Gaps 29;
OY 137 REQYLELSQUTHSITEIGPYTLDRDLSLYNGFTHRSSVPTTSQVSE-----PTLN 191

```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: October 14, 2006, 04:04:35 ; Search time 54 Seconds
(without alignments)
1212.461 Million cell updates/sec

Title: US-10-687-035-1

Perfect score: 3945

Sequence: 1 AAQPARARTRKLFTHRSV.....QKLISEDLNMTGHHHHH 748

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 650591 seqs, 87530628 residues

Total number of hits satisfying chosen parameters: 650591

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Issued Patents AA:*

- 1: /EMC_Celerra_SIDS3/prodata/2/iaa/5_COMB.pep:*
- 2: /EMC_Celerra_SIDS3/prodata/2/iaa/6_COMB.pep:*
- 3: /EMC_Celerra_SIDS3/prodata/2/iaa/7_COMB.pep:*
- 4: /EMC_Celerra_SIDS3/prodata/2/iaa/H_COMB.pep:*
- 5: /EMC_Celerra_SIDS3/prodata/2/iaa/CTUS_COMB.pep:*
- 6: /EMC_Celerra_SIDS3/prodata/2/iaa/RE_COMB.pep:*
- 7: /EMC_Celerra_SIDS3/prodata/2/iaa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|--------------------|
| 1 | 3663 | 92.9 | 833 | 2 | US-09-404-879A-389 |
| 2 | 3663 | 92.9 | 833 | 2 | US-09-667-857-389 |
| 3 | 3663 | 92.9 | 833 | 2 | US-10-198-053-389 |
| 4 | 3663 | 92.9 | 833 | 2 | US-09-827-271-389 |
| 5 | 3663 | 92.9 | 914 | 2 | US-09-404-879A-312 |
| 6 | 3663 | 92.9 | 914 | 2 | US-09-338-933-312 |
| 7 | 3663 | 92.9 | 914 | 2 | US-09-667-857-312 |
| 8 | 3663 | 92.9 | 914 | 2 | US-10-198-053-312 |
| 9 | 3663 | 92.9 | 914 | 2 | US-10-198-053-312 |
| 10 | 3663 | 92.9 | 914 | 2 | US-09-827-271-312 |
| 11 | 3651 | 92.5 | 3451 | 2 | US-10-198-053-595 |
| 12 | 3642 | 92.3 | 1148 | 2 | US-10-198-053-458 |
| 13 | 3642 | 92.3 | 1148 | 2 | US-10-198-053-478 |
| 14 | 3642 | 92.3 | 1148 | 2 | US-09-827-271-458 |
| 15 | 3642 | 92.3 | 1156 | 2 | US-10-198-053-459 |
| 16 | 3642 | 92.3 | 1156 | 2 | US-09-827-271-459 |
| 17 | 3189 | 80.8 | 772 | 2 | US-09-404-879A-388 |
| 18 | 3189 | 80.8 | 772 | 2 | US-09-667-857-388 |
| 19 | 3189 | 80.8 | 772 | 2 | US-10-198-053-388 |
| 20 | 3189 | 80.8 | 772 | 2 | US-09-827-271-388 |
| 21 | 1996 | 50.6 | 438 | 2 | US-10-198-053-483 |
| 22 | 1990 | 50.4 | 438 | 2 | US-09-404-879A-390 |
| 23 | 1990 | 50.4 | 438 | 2 | US-09-667-857-390 |
| 24 | 1990 | 50.4 | 438 | 2 | US-10-198-053-390 |
| 25 | 1990 | 50.4 | 438 | 2 | US-09-827-271-390 |
| 26 | 1436 | 36.4 | 304 | 2 | US-10-198-053-486 |

| | | | | | | |
|----|-------|------|-----|---|-------------------|-------------------|
| 27 | 1345 | 34.1 | 318 | 2 | US-10-198-053-594 | Sequence 594, App |
| 28 | 903 | 22.9 | 178 | 2 | US-10-198-053-489 | Sequence 489, App |
| 29 | 903 | 22.9 | 233 | 2 | US-10-198-053-488 | Sequence 488, App |
| 30 | 842.5 | 21.4 | 294 | 2 | US-10-198-053-487 | Sequence 487, App |
| 31 | 814 | 20.6 | 313 | 2 | US-10-198-053-461 | Sequence 461, App |
| 32 | 814 | 20.6 | 313 | 2 | US-09-827-271-461 | Sequence 461, App |
| 33 | 809 | 20.5 | 155 | 2 | US-10-198-053-591 | Sequence 591, App |
| 34 | 800 | 20.3 | 150 | 2 | US-10-198-053-593 | Sequence 593, App |
| 35 | 697 | 17.7 | 134 | 2 | US-10-198-053-592 | Sequence 592, App |
| 36 | 675.5 | 17.1 | 210 | 2 | US-10-198-053-481 | Sequence 481, App |
| 37 | 662 | 16.8 | 230 | 2 | US-10-198-053-480 | Sequence 480, App |
| 38 | 659.5 | 16.7 | 156 | 2 | US-10-198-053-589 | Sequence 589, App |
| 39 | 651.5 | 16.5 | 156 | 2 | US-10-198-053-596 | Sequence 596, App |
| 40 | 637.5 | 16.2 | 156 | 2 | US-10-198-053-590 | Sequence 590, App |
| 41 | 611.5 | 15.5 | 268 | 2 | US-10-198-053-485 | Sequence 485, App |
| 42 | 607.5 | 15.4 | 156 | 2 | US-10-198-053-584 | Sequence 584, App |
| 43 | 600.5 | 15.2 | 156 | 2 | US-10-198-053-587 | Sequence 587, App |
| 44 | 599.5 | 15.2 | 156 | 2 | US-10-198-053-583 | Sequence 583, App |
| 45 | 598.5 | 15.2 | 156 | 2 | US-10-198-053-586 | Sequence 586, App |

ALIGNMENTS

RESULT 1

US-09-404-879A-389
Sequence 389, Application US/09404879A
Patent No. 6468546
GENERAL INFORMATION:
APPLICANT: Mitcham, Jennifer L.
APPLICANT: King, Gordon E.
TITLE OF INVENTION: ALGATE, PAUL A.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.462C2
CURRENT APPLICATION NUMBER: US/09/404,879A
CURRENT FILING DATE: 1999-09-24
NUMBER OF SEQ ID NOS: 393
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 389
LENGTH: 833
TYPE: PRT
ORGANISM: Homo sapiens
US-09-404-879A-389

Query Match

Best Local Similarity 92.9%; Score 3663; DB 2; Length 833;
Matches 696; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

| | | |
|----|-----|---|
| QY | 14 | FTHRSSVSTTTPGPTVYLGAASKTPASIFGPAASHLLILFTLNTTNLRYEENWPG 73 |
| DB | 78 | FTHRSSVSTTTPGPTVYLGAASKTPASIFGPAASHLLILFTLNTTNLRYEENWPG 137 |
| QY | 74 | SRKENTTEVLGLRLPLFKNTSVGLYSGCRLLTRPKDGRATVDALCTHRPPTGP 133 |
| DB | 138 | SRKENTTEVLGLRLPLFKNTSVGLYSGCRLLTRPKDGRATVDALCTHRPPTGP 197 |
| QY | 134 | GLDREOLVELSQTLSITELGPTLDRDSLVLVNGFTHRSSVPTTSTGVSEEPFLNFT 193 |
| DB | 198 | GLDREOLVELSQTLSITELGPTLDRDSLVLVNGFTHRSSVPTTSTGVSEEPFLNFT 257 |
| QY | 194 | INNLRYMADMGPGSLIKENITDNVMKHLISPLFORSSLCARYTGCVIALRSVNGAETR 253 |
| DB | 258 | INNLRYMADMGPGSLIKENITDNVMKHLISPLFORSSLCARYTGCVIALRSVNGAETR 317 |
| QY | 254 | VLLLCYVLOPLSGPGPIKOVFHELSQOHTGTRIGPVSLLDQSLVLNKNBGPDEPT 313 |
| DB | 318 | VLLLCYVLOPLSGPGPIKOVFHELSQOHTGTRIGPVSLLDQSLVLNKNBGPDEPT 377 |
| QY | 314 | TRKPAATFPPLSEATTANGYHLKTLTLNFTISNLQYSPDMGKSGATFNSBEGVLOHLAR 373 |
| DB | 378 | TRKPAATFPPLSEATTANGYHLKTLTLNFTISNLQYSPDMGKSGATFNSBEGVLOHLAR 437 |

QY 374 PLFOKSSMGPPYLGQQLSLRPEKGAATGVDTCTHPPDPVGPGLDIQQLYWEISQULTH 433
DB 438 PLFOKSSMGPPYLGQQLSLRPEKGAATGVDTCTHPPDPVGPGLDIQQLYWEISQULTH 497
QY 434 GVTQGFVYLDRLDSLFINGYAPQNLISGEYQINFIHIVNMNLSPDPSSSEYITLLRDIO 493
DB 498 GVTQGFVYLDRLDSLFINGYAPQNLISGEYQINFIHIVNMNLSPDPSSSEYITLLRDIO 557
QY 494 DKVTLLYKGSQULHDFRCLVNLTMDSVLYTVKALFSSNLDPSLVEQVFLDKTLNASFH 553
DB 558 DKVTLLYKGSQULHDFRCLVNLTMDSVLYTVKALFSSNLDPSLVEQVFLDKTLNASFH 617
QY 554 WIGSTYQVLDIHTVEMESSVYQPTSSSTQHFYLNFTITNLPSQDKAQPGTTYNQRNR 613
DB 618 WIGSTYQVLDIHTVEMESSVYQPTSSSTQHFYLNFTITNLPSQDKAQPGTTYNQRNR 677
QY 614 NIEDALNOLFNRSSIKSYFSDCQVSTFRSVPNRHHTGVDSLGNFSPILARVDVAIYEEF 673
DB 678 NIEDALNOLFNRSSIKSYFSDCQVSTFRSVPNRHHTGVDSLGNFSPILARVDVAIYEEF 737
QY 674 LAMTRNGTQLONFITLDRSSVLYDGYSPNRNEPLTGN 710
DB 738 LAMTRNGTQLONFITLDRSSVLYDGYSPNRNEPLTGN 774

RESULT 2
US-09-667-857-389
Sequence 389, Application US/09667857
Patent No. 669664
GENERAL INFORMATION:
APPLICANT: Mitcham, Jennifer L.
APPLICANT: King, Gordon E.
APPLICANT: Algate, Paul A.
APPLICANT: Fling, Steven P.
APPLICANT: Retter, Marc W.
APPLICANT: Fanger, Gary Richard
APPLICANT: Reed, Steven G.
APPLICANT: Vedvick, Thomas S.
APPLICANT: Carter, Darick
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF OVARIAN CANCER
FILE REFERENCE: 210121.462C5
CURRENT APPLICATION NUMBER: US/09/667,857
CURRENT FILING DATE: 2000-09-20
NUMBER OF SEQ ID NOS: 455
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 389
LENGTH: 833
TYPE: PRF
ORGANISM: Homo sapiens
US-09-667-857-389

Query Match 92.9%; Score 3663; DB 2; Length 833;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 696; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 14 FTHRSSVSTSTGTPTVYLGASKTTPASIFGSAASHLLILFTLNTITNLRYEENMWP 73
DB 78 FTHRSSVSTSTGTPTVYLGASKTTPASIFGSAASHLLILFTLNTITNLRYEENMWP 137
QY 74 SRKFNTERVLOGLRPLFRNTSVGLYSGCRLLTLRPEKGAATGVDAICTHRPDTGP 133
DB 138 SRKFNTERVLOGLRPLFRNTSVGLYSGCRLLTLRPEKGAATGVDAICTHRPDTGP 197
QY 134 GLDRBOLYLELSQULTHSITELGPTLDRSLVYNGFTHRSSVPTSTGVVSEEPFTLNFT 193
DB 198 GLDRBOLYLELSQULTHSITELGPTLDRSLVYNGFTHRSSVPTSTGVVSEEPFTLNFT 257
QY 194 INNLRYADMGQPSLKFNTITDNVMKHLSPLEFORSISGARVYTCRVIARSVNGAETR 253
DB 258 INNLRYADMGQPSLKFNTITDNVMKHLSPLEFORSISGARVYTCRVIARSVNGAETR 317
QY 254 VDLCTYQLPISGGLPIKOVFHELSQOHTGTRLGYSLDKDSL YLANGVNEPDPDEPPT 313

DB 318 VDLCTYQLPISGGLPIKOVFHELSQOHTGTRLGYSLDKDSL YLANGVNEPDPDEPPT 377
QY 314 TPKPATITLPLSEATTTMGMHLKTLTNTFTISNLQVSPDMKGSATNSGTVQHLR 373
DB 378 TPKPATITLPLSEATTTMGMHLKTLTNTFTISNLQVSPDMKGSATNSGTVQHLR 437
QY 374 PLFOKSSMGPPYLGQQLSLRPEKGAATGVDTCTHPPDPVGPGLDIQQLYWEISQULTH 433
DB 438 PLFOKSSMGPPYLGQQLSLRPEKGAATGVDTCTHPPDPVGPGLDIQQLYWEISQULTH 497
QY 434 GVTQGFVYLDRLDSLFINGYAPQNLISGEYQINFIHIVNMNLSPDPSSSEYITLLRDIO 493
DB 498 GVTQGFVYLDRLDSLFINGYAPQNLISGEYQINFIHIVNMNLSPDPSSSEYITLLRDIO 557
QY 494 DKVTLLYKGSQULHDFRCLVNLTMDSVLYTVKALFSSNLDPSLVEQVFLDKTLNASFH 553
DB 558 DKVTLLYKGSQULHDFRCLVNLTMDSVLYTVKALFSSNLDPSLVEQVFLDKTLNASFH 617
QY 554 WIGSTYQVLDIHTVEMESSVYQPTSSSTQHFYLNFTITNLPSQDKAQPGTTYNQRNR 613
DB 618 WIGSTYQVLDIHTVEMESSVYQPTSSSTQHFYLNFTITNLPSQDKAQPGTTYNQRNR 677
QY 614 NIEDALNOLFNRSSIKSYFSDCQVSTFRSVPNRHHTGVDSLGNFSPILARVDVAIYEEF 673
DB 678 NIEDALNOLFNRSSIKSYFSDCQVSTFRSVPNRHHTGVDSLGNFSPILARVDVAIYEEF 737
QY 674 LAMTRNGTQLONFITLDRSSVLYDGYSPNRNEPLTGN 710
DB 738 LAMTRNGTQLONFITLDRSSVLYDGYSPNRNEPLTGN 774

RESULT 3
US-10-198-053-389
Sequence 389, Application US/10198053
Patent No. 6858710
GENERAL INFORMATION:
APPLICANT: Bangur, Chaitanya S.
APPLICANT: Retter, Marc W.
APPLICANT: Fanger, Gary R.
APPLICANT: Hill, Paul
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
FILE REFERENCE: 210121.462C9
CURRENT APPLICATION NUMBER: US/10/198,053
CURRENT FILING DATE: 2002-07-17
NUMBER OF SEQ ID NOS: 624
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 389
LENGTH: 833
TYPE: PRF
ORGANISM: Homo sapiens
US-10-198-053-389

Query Match 92.9%; Score 3663; DB 2; Length 833;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 696; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 14 FTHRSSVSTSTGTPTVYLGASKTTPASIFGSAASHLLILFTLNTITNLRYEENMWP 73
DB 78 FTHRSSVSTSTGTPTVYLGASKTTPASIFGSAASHLLILFTLNTITNLRYEENMWP 137
QY 74 SRKFNTERVLOGLRPLFRNTSVGLYSGCRLLTLRPEKGAATGVDAICTHRPDTGP 133
DB 138 SRKFNTERVLOGLRPLFRNTSVGLYSGCRLLTLRPEKGAATGVDAICTHRPDTGP 197
QY 134 GLDRBOLYLELSQULTHSITELGPTLDRSLVYNGFTHRSSVPTSTGVVSEEPFTLNFT 193
DB 198 GLDRBOLYLELSQULTHSITELGPTLDRSLVYNGFTHRSSVPTSTGVVSEEPFTLNFT 257
QY 194 INNLRYADMGQPSLKFNTITDNVMKHLSPLEFORSISGARVYTCRVIARSVNGAETR 253
DB 258 INNLRYADMGQPSLKFNTITDNVMKHLSPLEFORSISGARVYTCRVIARSVNGAETR 317

QY 254 VDLICTYLOPLSGRGLPIKOVFHEL.SOOTHTGTRIGPYSLDKOSLYLINGNEPDPDEPPT 313
DB 318 VDLICTYLOPLSGRGLPIKOVFHEL.SOOTHTGTRIGPYSLDKOSLYLINGNEPDPDEPPT 377
QY 314 TPKPATTFLPPLSEATTAMGYHLKTLINFTISNLQYSPDMGKSATFNSTEGVLOHLR 373
DB 378 TPKPATTFLPPLSEATTAMGYHLKTLINFTISNLQYSPDMGKSATFNSTEGVLOHLR 437
QY 374 PLFOKSMGPFYIGCOLISLRPEKGAATGVDTTCTYHPDPVPGGLDIOQLYWELSQLTH 433
DB 438 PLFOKSMGPFYIGCOLISLRPEKGAATGVDTTCTYHPDPVPGGLDIOQLYWELSQLTH 497
QY 434 GTVQLGFYVLDROSLFINGAPONLSIRGEYQINFIHVMNLSNPDPSTSEYITLLDIO 493
DB 498 GTVQLGFYVLDROSLFINGAPONLSIRGEYQINFIHVMNLSNPDPSTSEYITLLDIO 557
QY 494 DKYTTLYKGSQDLHDFRCLVTNLTMDSVLYVYKALFSSNLDPSELVEQVFLDKTLNASFH 553
DB 558 DKYTTLYKGSQDLHDFRCLVTNLTMDSVLYVYKALFSSNLDPSELVEQVFLDKTLNASFH 617
QY 554 WLGSTYQVLDIHTEMESSYVQPTSSSTQHFIYLNFTITNL.PYSQDKAQPGTTNYQRNKR 613
DB 618 WLGSTYQVLDIHTEMESSYVQPTSSSTQHFIYLNFTITNL.PYSQDKAQPGTTNYQRNKR 677
QY 614 NIEDANOLFRRNSISYSDCOVSTFRSVPNRHHTGVDSL.CNFSPLARRVDVAIYEER 673
DB 678 NIEDANOLFRRNSISYSDCOVSTFRSVPNRHHTGVDSL.CNFSPLARRVDVAIYEER 737
QY 674 LRMTNGTOLQNTFLDRSSVLYDGYSPNRNEPLTGN 710
DB 738 LRMTNGTOLQNTFLDRSSVLYDGYSPNRNEPLTGN 774

RESULT 4
US-09-827-271-389
; Sequence 389, Application US/09827271
; Patent No. 6962980
; GENERAL INFORMATION:
; APPLICANT: Retter, Marc W.
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.462C6
; CURRENT FILING DATE: 2001-04-04
; NUMBER OF SEQ ID NOS: 461
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 389
; LENGTH: 833
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-827-271-389

Query Match 92.9%; Score 3663; DB 2; Length 833;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 696; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 14 FTHRSSVSTSTGTPTVYLGASKTPASIFGSAASHLLILFTLNTITNLRYEENMMPG 73
DB 78 FTHRSSVSTSTGTPTVYLGASKTPASIFGSAASHLLILFTLNTITNLRYEENMMPG 137
QY 74 SRKFNTTERVLQGLRPLFKNTSVGPLYSGCRLLTLRPEKDGATGVDAICTHRPDPGP 133
DB 138 SRKFNTTERVLQGLRPLFKNTSVGPLYSGCRLLTLRPEKDGATGVDAICTHRPDPGP 197
QY 134 GLDREQLYLELSQTHSITELGPLYTLDRDSL.YVNGFTHRSSVPTTSTGVVSEEPFLNFT 133
DB 198 GLDREQLYLELSQTHSITELGPLYTLDRDSL.YVNGFTHRSSVPTTSTGVVSEEPFLNFT 257
QY 194 INNLRYAMDMGPGSLKFNITDVMKHLSP.LFORSSLGARYGCRVIALRSYKGAETR 253
DB 258 INNLRYAMDMGPGSLKFNITDVMKHLSP.LFORSSLGARYGCRVIALRSYKGAETR 317

QY 254 VDLICTYLOPLSGRGLPIKOVFHEL.SOOTHTGTRIGPYSLDKOSLYLINGNEPDPDEPPT 313
DB 318 VDLICTYLOPLSGRGLPIKOVFHEL.SOOTHTGTRIGPYSLDKOSLYLINGNEPDPDEPPT 377
QY 314 TPKPATTFLPPLSEATTAMGYHLKTLINFTISNLQYSPDMGKSATFNSTEGVLOHLR 373
DB 378 TPKPATTFLPPLSEATTAMGYHLKTLINFTISNLQYSPDMGKSATFNSTEGVLOHLR 437
QY 374 PLFOKSMGPFYIGCOLISLRPEKGAATGVDTTCTYHPDPVPGGLDIOQLYWELSQLTH 433
DB 438 PLFOKSMGPFYIGCOLISLRPEKGAATGVDTTCTYHPDPVPGGLDIOQLYWELSQLTH 497
QY 434 GTVQLGFYVLDROSLFINGAPONLSIRGEYQINFIHVMNLSNPDPSTSEYITLLDIO 493
DB 498 GTVQLGFYVLDROSLFINGAPONLSIRGEYQINFIHVMNLSNPDPSTSEYITLLDIO 557
QY 494 DKYTTLYKGSQDLHDFRCLVTNLTMDSVLYVYKALFSSNLDPSELVEQVFLDKTLNASFH 553
DB 558 DKYTTLYKGSQDLHDFRCLVTNLTMDSVLYVYKALFSSNLDPSELVEQVFLDKTLNASFH 617
QY 554 WLGSTYQVLDIHTEMESSYVQPTSSSTQHFIYLNFTITNL.PYSQDKAQPGTTNYQRNKR 613
DB 618 WLGSTYQVLDIHTEMESSYVQPTSSSTQHFIYLNFTITNL.PYSQDKAQPGTTNYQRNKR 677
QY 614 NIEDANOLFRRNSISYSDCOVSTFRSVPNRHHTGVDSL.CNFSPLARRVDVAIYEER 673
DB 678 NIEDANOLFRRNSISYSDCOVSTFRSVPNRHHTGVDSL.CNFSPLARRVDVAIYEER 737
QY 674 LRMTNGTOLQNTFLDRSSVLYDGYSPNRNEPLTGN 710
DB 738 LRMTNGTOLQNTFLDRSSVLYDGYSPNRNEPLTGN 774

RESULT 5
US-09-404-879A-312
; Sequence 312, Application US/09404879A
; Patent No. 6468546
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; APPLICANT: Algate, Paul A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.462C2
; CURRENT FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 393
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 312
; LENGTH: 914
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-404-879A-312

Query Match 92.9%; Score 3663; DB 2; Length 914;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 696; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 14 FTHRSSVSTSTGTPTVYLGASKTPASIFGSAASHLLILFTLNTITNLRYEENMMPG 73
DB 159 FTHRSSVSTSTGTPTVYLGASKTPASIFGSAASHLLILFTLNTITNLRYEENMMPG 218
QY 74 SRKFNTTERVLQGLRPLFKNTSVGPLYSGCRLLTLRPEKDGATGVDAICTHRPDPGP 133
DB 219 SRKFNTTERVLQGLRPLFKNTSVGPLYSGCRLLTLRPEKDGATGVDAICTHRPDPGP 278
QY 134 GLDREQLYLELSQTHSITELGPLYTLDRDSL.YVNGFTHRSSVPTTSTGVVSEEPFLNFT 193
DB 279 GLDREQLYLELSQTHSITELGPLYTLDRDSL.YVNGFTHRSSVPTTSTGVVSEEPFLNFT 338
QY 194 INNLRYAMDMGPGSLKFNITDVMKHLSP.LFORSSLGARYGCRVIALRSYKGAETR 253
DB 258 INNLRYAMDMGPGSLKFNITDVMKHLSP.LFORSSLGARYGCRVIALRSYKGAETR 317

| | | | |
|----|-----|---|-----|
| Db | 339 | INNLRYMADMGPGSLKFRNITDNWAKHLISPLFOSSSLGARITGCIVIALRSTKQNAEIR | 398 |
| Oy | 254 | VDLCTLYLOPLSGPGPLIKQVFEHLSQOQTHTGTRGLRPSGLDKDSLVLNGNEBGPDEPT | 313 |
| Db | 399 | VDLCTLYLOPLSGPGPLIKQVFEHLSQOQTHTGTRGLRPSGLDKDSLVLNGNEBGPDEPT | 458 |
| Oy | 314 | TPKATPTPLPRLSBAITTAGYHLKLTLLTLPFTLSNLOYSBDMGKGSATPNSBEBVLHLLR | 373 |
| Db | 459 | TPKATPTPLPRLSBAITTAGYHLKLTLLTLPFTLSNLOYSBDMGKGSATPNSBEBVLHLLR | 518 |
| Oy | 374 | PLFOKSSMGPFLYGCCLISLRPEKDAAGTVDITTCYHDPVPGSLDIQOLYWELSQLT | 433 |
| Db | 519 | PLFOKSSMGPFLYGCCLISLRPEKDAAGTVDITTCYHDPVPGSLDIQOLYWELSQLT | 578 |
| Oy | 434 | GVTQOLGFYVLDBDLSFLINGVAPONLSIRBEOYNFIHVMNLSNMPDTSSEYITLLRDIO | 493 |
| Db | 579 | GVTQOLGFYVLDBDLSFLINGVAPONLSIRBEOYNFIHVMNLSNMPDTSSEYITLLRDIO | 638 |
| Oy | 494 | DKVITLLYKKSQHLHDFRFLVTNLITMDSVLVYKALFSSNLDBSLVEQVFLDKTLNASF | 553 |
| Db | 639 | DKVITLLYKKSQHLHDFRFLVTNLITMDSVLVYKALFSSNLDBSLVEQVFLDKTLNASF | 698 |
| Oy | 554 | WLGSTYQVLVDIHYTEMESSYVQPTSSSSSTOCHFNLNFTITTLPRYSQDAQOAGTNYQNK | 613 |
| Db | 699 | WLGSTYQVLVDIHYTEMESSYVQPTSSSSSTOCHFNLNFTITTLPRYSQDAQOAGTNYQNK | 758 |
| Oy | 614 | NIEDALNOLFRRNSIKSYFSDCQVSTFRSPVNNHHTGVDSLCLNPSPLARRVDRAIYEEF | 673 |
| Db | 759 | NIEDALNOLFRRNSIKSYFSDCQVSTFRSPVNNHHTGVDSLCLNPSPLARRVDRAIYEEF | 818 |
| Oy | 674 | LKMTNGTQLOLNTLLDBSSVLVDGYSPNRRNEPLTGN | 710 |
| Db | 819 | LKMTNGTQLOLNTLLDBSSVLVDGYSPNRRNEPLTGN | 855 |

| | | |
|-----|---|-----|
| 339 | INLRMYMADMGQPGSKLFENITDYNMKLLBSLFQBRSSJGARYTTCGRIALRVSXNQAETR | 398 |
| OY | VDLLCTYLOPLSGPGPLIKOVFHELSQOITHG1TRIGPYSLDKOSLYLNGYNEPQDEPPT | 313 |
| Db | VDLLCTYLOPLSGPGPLIKOVFHELSQOITHG1TRIGPYSLDKOSLYLNGYNEPQDEPPT | 458 |
| OY | TPRPATTFPLPSEATTAAGYHLKTLTLNFTISNLQVSPDMGKSATFNS7BEGVLOHLR | 373 |
| Db | TPRPATTFPLPSEATTAAGYHLKTLTLNFTISNLQVSPDMGKSATFNS7BEGVLOHLR | 518 |
| OY | PLFOKSMGPFYLCOLISLRPEKDGAA7GVDTTCYHPDPVPGQLDIQOLYMLSOULTH | 433 |
| Db | PLFOKSMGPFYLCQGLISLRPEKDGAA7GVDTTCYHPDPVPGQLDIQOLYMLSOULTH | 578 |
| OY | GV7QLGFFYLDRBSLFINGTAPONIS7IGAEOVINHIYNNMLSNPDP7SSEYITLLADIQ | 493 |
| Db | GV7QLGFFYLDRBSLFINGTAPONIS7IKEEYOINHIYNNMLSNPDP7SSEYITLLADIQ | 638 |
| OY | DKYTLTYKSGSLHD7RFCLVNTLTMDSVLVTVKALFSSNLDPSLVEQVFLDKTLNASFH | 555 |
| Db | DKYTLTYKSGSLHD7RFCLVNTLTMDSVLVTVKALFSSNLDPSLVEQVFLDKTLNASFH | 698 |
| OY | WLGSTYLOVDIHTBNESSY7OPTSSSS7QHLYNLFTITNL7PYSDDKAQOECTNYOANKR | 613 |
| Db | WLGSTYLOVDIHTBNESSY7OPTSSSS7QHLYNLFTITNL7PYSDDKAQOECTNYOANKR | 758 |
| OY | NIEBALNQLFRNSSISYSFSDCOVS7FRSVPRHHHTGVDSLCNFSPILARVDRAVAIYEEF | 673 |
| Db | NIEBALNQLFRNSSISYSFSDCOVS7FRSVPRHHHTGVDSLCNFSPILARVDRAVAIYEEF | 818 |
| OY | L7RMTNGTQLONTFLDRSSVYLVDGVS7PNNNEP7LGN5 | 710 |
| Db | L7RMTNGTQLONTFLDRSSVYLVDGVS7PNNNEP7LGN5 | 855 |

```

RESULT 6
US-09-338-933-312
/ Sequence 312, Application US/09338933
/ Patent No. 6488931
/ GENERAL INFORMATION:
/ APPLICANT: Mitcham, Jennifer Lynn
/ APPLICANT: King, Gordon B.
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY OF
/ TITLE OF INVENTION: OVARIAN CANCER
/ FILE REFERENCE: 210121.462C1
/ CURRENT APPLICATION NUMBER: US/09/338.933
/ CURRENT FILING DATE: 1999-06-23
/ NUMBER OF SEQ ID NOS: 312
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 312
/ LENGTH: 914
/ TYPE: PRT
/ ORGANISM: Homo sapien
US-09-338-933-312

```

| | | | | |
|---------------------------|-------|---------------|-----------|-------------|
| Query Match | 92.9% | Score 3663; | DB 2; | Length 914; |
| Best Local Similarity | 99.9% | Pred. No. 0; | | |
| Matches 696; Conservative | 0; | Mismatches 1; | Indels 0; | Gaps 0; |

| | | | |
|----|-----|--|-----|
| Qy | 14 | FTHRSSVSTSTGPTVIVGASKTPASIFGSSASHLLILTLNLTINILYEENMMPG | 73 |
| Db | 159 | FTHRSSVSTSTGPTVIVGASKTPASIFGSSASHLLILTLNLTINILYEENMMPG | 218 |
| Qy | 74 | SRKFNTERVLOGLRPLFNKTSVGPLYSGCRLLTLRPEKDBAGTVDALCTHRDPPTGP | 133 |
| Db | 219 | SRKFNTERVLOGLRPLFNKTSVGPLYSGCRLLTLRPEKDBAGTVDALCTHRDPPTGP | 278 |
| Qy | 134 | GLDRBOUYELLSQUTSRITELGPPYTLDRSLYNGFTHRSSVPTTSTGVVSEEPFLNLT | 193 |
| Db | 279 | GLDRBOUYELLSQUTSRITELGPPYTLDRSLYNGFTHRSSVPTTSTGVVSEEPFLNLT | 338 |
| Qy | 194 | INMLRYADMGQPGSLKFNITIDNMKHLSPLEFORSSLAGRYTGCRVIALRSVKGAEFR | 253 |

```

RESULT 7
US-09-667-857-312
; Sequence 312, Application US/09667857
; Patent No. 6699664
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; APPLICANT: Algate, Paul A.
; APPLICANT: Fling, Steven P.
; APPLICANT: Reltter, Marc W.
; APPLICANT: Fanger, Gary Richard
; APPLICANT: Reed, Steven G.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCES: 210121.462C5
; CURRENT APPLICATION NUMBER: US/09/667,857
; CURRENT FILING DATE: 2000-09-20
; NUMBER OF SEQ ID NOS: 455
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 312
; LENGTH: 914
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-667-857-312

```

| | | | | |
|---------------------------|--------|---------------|-----------|-------------|
| Query Match | 92.9% | Score 3663; | DB 2; | Length 914; |
| Best Local Similarity | 99.9%; | Pred. No. 0; | | |
| Matches 696; Conservative | 0; | Mismatches 1; | Indels 0; | Gaps 0; |

| | | | | | | | | |
|----|-----|------------------------|--------|------------|--------------|----------|----|-----|
| Qy | 14 | FTHRSSVSTSTSPGPPTVYLGA | SKTPAS | FGSASASHLL | LTPLNFTPLN | RYEENM | PG | 73 |
| Db | 159 | FTHRSSVSTSTSPGPPTVYLGA | SKTPAS | FGSASASHLL | LTPLNFTPLN | RYEENM | PG | 216 |
| Qy | 74 | SRKNTTERTVLOGLLRPLFKNT | SVGL | YSGCRLTL | LRPKCGEANGVD | LTCHRPDP | PG | 133 |
| Db | 219 | SRKNTTERTVLOGLLRPLFKNT | SVGL | YSGCRLTL | LRPKCGEANGVD | LTCHRPDP | PG | 278 |

QY 134 GLDREQLYELSQLTHSITELGPLYLDRDSL.YVNGFTHRSSVPTTSTGVSEEPFTLNFT 193
DB 279 GLDREQLYELSQLTHSITELGPLYLDRDSL.YVNGFTHRSSVPTTSTGVSEEPFTLNFT 338
QY 194 INNLRYADMGQPSLSKFNTIDNMKHLSPFORSSIGARYTCRYIALRSVNGAETR 253
DB 339 INNLRYADMGQPSLSKFNTIDNMKHLSPFORSSIGARYTCRYIALRSVNGAETR 398
QY 254 VDLICTYLOPLSGGLPIKOVFHELSQOQTHGIRLGRYSIDKDSL.YVNGNEGPDPPT 313
DB 399 VDLICTYLOPLSGGLPIKOVFHELSQOQTHGIRLGRYSIDKDSL.YVNGNEGPDPPT 458
QY 314 TPKPATTFLPLSBATTAMGYHLKTLTNFTISNLOYSPDMGKSATFNSTEGVLOHLR 373
DB 459 TPKPATTFLPLSBATTAMGYHLKTLTNFTISNLOYSPDMGKSATFNSTEGVLOHLR 518
QY 374 PLFOKSMGPFYLGCOLISLRPEKGAATGVDTTCTYHPDPVPGGLDIQOLYWELSQUTH 433
DB 519 PLFOKSMGPFYLGCOLISLRPEKGAATGVDTTCTYHPDPVPGGLDIQOLYWELSQUTH 578
QY 434 GVTOLGFYVLDRLSLFINGYAPQNLISRGEXQINPHI.VNMNLSNPDPTSEYITLLRDIQ 493
DB 579 GVTOLGFYVLDRLSLFINGYAPQNLISRGEXQINPHI.VNMNLSNPDPTSEYITLLRDIQ 638
QY 494 DKYTTLYKGSQLDHTRFCLVTNL.TMDSVLYTVKAL.FSSNLDPSLVEQVFLDKTLNLSFH 553
DB 639 DKYTTLYKGSQLDHTRFCLVTNL.TMDSVLYTVKAL.FSSNLDPSLVEQVFLDKTLNLSFH 698
QY 554 WLGSTYOLVDIHTEMESSEYVQPTSSSTQHFYLNFTTNLPYSQDAQOCTTYOQNR 613
DB 699 WLGSTYOLVDIHTEMESSEYVQPTSSSTQHFYLNFTTNLPYSQDAQOCTTYOQNR 758
QY 614 NIEDALNQLFRNSSIKSYFSDCOVSTFRSVPNRHHTGVDSL.CNFSPLARVDVAIYEEF 673
DB 759 NIEDALNQLFRNSSIKSYFSDCOVSTFRSVPNRHHTGVDSL.CNFSPLARVDVAIYEEF 818
QY 674 LRMTNGTOLQONFTLDRSSVLYVDGYSFNRNEPPLTGN 710
DB 819 LRMTNGTOLQONFTLDRSSVLYVDGYSFNRNEPPLTGN 855

RESULT 8
US-10-198-053-312
; Sequence 312, Application US/10198053
; Patent No. 6858710
; GENERAL INFORMATION:
; APPLICANT: Bangur, Chaityanya S.
; APPLICANT: Rafter, Marc W.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Hill, Paul
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.462C9
; CURRENT APPLICATION NUMBER: US/10/198.053
; CURRENT FILING DATE: 2002-07-17
; NUMBER OF SEQ ID NOS: 624
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 312
; LENGTH: 914
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-198-053-312

Query Match 92.9%; Score 3663; DB 2; Length 914;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 696; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 14 FTHRSSVSTSTGTPTVYLGASKTSPASIFGSAASHLLILFTLNFTTNLYEENMMPG 73
DB 159 FTHRSSVSTSTGTPTVYLGASKTSPASIFGSAASHLLILFTLNFTTNLYEENMMPG 218
QY 74 SRKNTTERVUQGLRPLFKNTSVGPLXSGCRLLTLPKEDGATGVDAICTHRPDPGP 133

DB 219 SRKNTTERVUQGLRPLFKNTSVGPLXSGCRLLTLPKEDGATGVDAICTHRPDPGP 278
QY 134 GLDREQLYELSQLTHSITELGPLYLDRDSL.YVNGFTHRSSVPTTSTGVSEEPFTLNFT 193
DB 279 GLDREQLYELSQLTHSITELGPLYLDRDSL.YVNGFTHRSSVPTTSTGVSEEPFTLNFT 338
QY 194 INNLRYADMGQPSLSKFNTIDNMKHLSPFORSSIGARYTCRYIALRSVNGAETR 253
DB 339 INNLRYADMGQPSLSKFNTIDNMKHLSPFORSSIGARYTCRYIALRSVNGAETR 398
QY 254 VDLICTYLOPLSGGLPIKOVFHELSQOQTHGIRLGRYSIDKDSL.YVNGNEGPDPPT 313
DB 399 VDLICTYLOPLSGGLPIKOVFHELSQOQTHGIRLGRYSIDKDSL.YVNGNEGPDPPT 458
QY 314 TPKPATTFLPLSBATTAMGYHLKTLTNFTISNLOYSPDMGKSATFNSTEGVLOHLR 373
DB 459 TPKPATTFLPLSBATTAMGYHLKTLTNFTISNLOYSPDMGKSATFNSTEGVLOHLR 518
QY 374 PLFOKSMGPFYLGCOLISLRPEKGAATGVDTTCTYHPDPVPGGLDIQOLYWELSQUTH 433
DB 519 PLFOKSMGPFYLGCOLISLRPEKGAATGVDTTCTYHPDPVPGGLDIQOLYWELSQUTH 578
QY 434 GVTOLGFYVLDRLSLFINGYAPQNLISRGEXQINPHI.VNMNLSNPDPTSEYITLLRDIQ 493
DB 579 GVTOLGFYVLDRLSLFINGYAPQNLISRGEXQINPHI.VNMNLSNPDPTSEYITLLRDIQ 638
QY 494 DKYTTLYKGSQLDHTRFCLVTNL.TMDSVLYTVKAL.FSSNLDPSLVEQVFLDKTLNLSFH 553
DB 639 DKYTTLYKGSQLDHTRFCLVTNL.TMDSVLYTVKAL.FSSNLDPSLVEQVFLDKTLNLSFH 698
QY 554 WLGSTYOLVDIHTEMESSEYVQPTSSSTQHFYLNFTTNLPYSQDAQOCTTYOQNR 613
DB 699 WLGSTYOLVDIHTEMESSEYVQPTSSSTQHFYLNFTTNLPYSQDAQOCTTYOQNR 758
QY 614 NIEDALNQLFRNSSIKSYFSDCOVSTFRSVPNRHHTGVDSL.CNFSPLARVDVAIYEEF 673
DB 759 NIEDALNQLFRNSSIKSYFSDCOVSTFRSVPNRHHTGVDSL.CNFSPLARVDVAIYEEF 818
QY 674 LRMTNGTOLQONFTLDRSSVLYVDGYSFNRNEPPLTGN 710
DB 819 LRMTNGTOLQONFTLDRSSVLYVDGYSFNRNEPPLTGN 855

RESULT 9
US-10-198-053-478
; Sequence 478, Application US/10198053
; Patent No. 6858710
; GENERAL INFORMATION:
; APPLICANT: Bangur, Chaityanya S.
; APPLICANT: Rafter, Marc W.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Hill, Paul
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.462C9
; CURRENT APPLICATION NUMBER: US/10/198.053
; CURRENT FILING DATE: 2002-07-17
; NUMBER OF SEQ ID NOS: 624
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 478
; LENGTH: 914
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-198-053-478

Query Match 92.9%; Score 3663; DB 2; Length 914;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 696; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 14 FTHRSSVSTSTGTPTVYLGASKTSPASIFGSAASHLLILFTLNFTTNLYEENMMPG 73
DB 159 FTHRSSVSTSTGTPTVYLGASKTSPASIFGSAASHLLILFTLNFTTNLYEENMMPG 218

QY 74 SRKFTTERTVLOGGLRPLFKNTSVGPLYSGCRLTLRPEKDEATGVDAICTHRDPDGP 133
DB 219 SRKFTTERTVLOGGLRPLFKNTSVGPLYSGCRLTLRPEKDEATGVDAICTHRDPDGP 278
QY 134 GLDRQVLYLELSQVTHSITELGPTLDRDSLVLVNGFTHRSSVPTTSTGVVSEEPFLNFT 193
DB 279 GLDRQVLYLELSQVTHSITELGPTLDRDSLVLVNGFTHRSSVPTTSTGVVSEEPFLNFT 338
QY 194 INNLRYMADMGQPGSLKFNITDVMKHLSPFORSSIGARTGCRVIALRSVKNGAETR 253
DB 339 INNLRYMADMGQPGSLKFNITDVMKHLSPFORSSIGARTGCRVIALRSVKNGAETR 398
QY 254 VDLICTYLOPLSGPGPIKQVHELSQOHTGTRIGPYSLDKDSLVLNGYNEPGDEPPT 313
DB 399 VDLICTYLOPLSGPGPIKQVHELSQOHTGTRIGPYSLDKDSLVLNGYNEPGDEPPT 458
QY 314 TPKPATTPLPPLSEATTAMGYHLKTLTNFTISNLQYSPDMKGSATFNSTEGVLOHLR 373
DB 459 TPKPATTPLPPLSEATTAMGYHLKTLTNFTISNLQYSPDMKGSATFNSTEGVLOHLR 518
QY 374 PLFOKSSMGPFYLGCOLISLRPEKDAATGVDTTCYHPDPVGPGLDIOQLYWELSQVTH 433
DB 519 PLFOKSSMGPFYLGCOLISLRPEKDAATGVDTTCYHPDPVGPGLDIOQLYWELSQVTH 578
QY 434 GVTOLGFYVLDRLSLFINGYAPONTLSIRGEYQINFHIVMNLSPDPTSSSEYITLLRDIO 493
DB 579 GVTOLGFYVLDRLSLFINGYAPONTLSIRGEYQINFHIVMNLSPDPTSSSEYITLLRDIO 638
QY 494 DKVTLLYKGSQVLDHTRFCLVNTLTMDSVLYTVKALFSSNLDPSLVEQVFLDKTLNASH 553
DB 639 DKVTLLYKGSQVLDHTRFCLVNTLTMDSVLYTVKALFSSNLDPSLVEQVFLDKTLNASH 698
QY 554 WLGSTYQVLDIHYTEMESVYQPTSSSTQHFYLNFTITNLPSYODKAPGTTNORNR 613
DB 699 WLGSTYQVLDIHYTEMESVYQPTSSSTQHFYLNFTITNLPSYODKAPGTTNORNR 758
QY 614 NIEDALNOLFRNSSISYSDCOVSTFRSVPNRHHTGVDSLGNFSLARVDRVAYIEEF 673
DB 759 NIEDALNOLFRNSSISYSDCOVSTFRSVPNRHHTGVDSLGNFSLARVDRVAYIEEF 818
QY 674 LRMTNRGTOLQNTFLDRSSVLYDGYSPNRNEPLTGN 710
DB 819 LRMTNRGTOLQNTFLDRSSVLYDGYSPNRNEPLTGN 855

RESULT 10
US-09-827-271-312
/ Sequence 312, Application US/09827271
/ Patent No. 6962980
/ GENERAL INFORMATION:
/ APPLICANT: Recter, Marc W.
/ APPLICANT: Recter, Gary R.
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
/ TITLE OF INVENTION: DIAGNOSIS OF OVARIAN CANCER
/ FILE REFERENCE: 210121.462C6
/ CURRENT APPLICATION NUMBER: US/09/827,271
/ CURRENT FILING DATE: 2001-04-04
/ NUMBER OF SEQ ID NOS: 461
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 312
/ LENGTH: 914
/ TYPE: PRT
/ ORGANISM: Homo sapien
US-09-827-271-312

Query Match 92.9%; Score 3663; DB 2; Length 914;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 696; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 14 FTHRSSVSTSTPTPTVYLGASTPASIFGSAASHLLIFLTNFTITNLRYEENMPG 73
DB 159 FTHRSSVSTSTPTPTVYLGASTPASIFGSAASHLLIFLTNFTITNLRYEENMPG 218

QY 74 SRKFTTERTVLOGGLRPLFKNTSVGPLYSGCRLTLRPEKDEATGVDAICTHRDPDGP 133
DB 219 SRKFTTERTVLOGGLRPLFKNTSVGPLYSGCRLTLRPEKDEATGVDAICTHRDPDGP 278
QY 134 GLDRQVLYLELSQVTHSITELGPTLDRDSLVLVNGFTHRSSVPTTSTGVVSEEPFLNFT 193
DB 279 GLDRQVLYLELSQVTHSITELGPTLDRDSLVLVNGFTHRSSVPTTSTGVVSEEPFLNFT 338
QY 194 INNLRYMADMGQPGSLKFNITDVMKHLSPFORSSIGARTGCRVIALRSVKNGAETR 253
DB 339 INNLRYMADMGQPGSLKFNITDVMKHLSPFORSSIGARTGCRVIALRSVKNGAETR 398
QY 254 VDLICTYLOPLSGPGPIKQVHELSQOHTGTRIGPYSLDKDSLVLNGYNEPGDEPPT 313
DB 399 VDLICTYLOPLSGPGPIKQVHELSQOHTGTRIGPYSLDKDSLVLNGYNEPGDEPPT 458
QY 314 TPKPATTPLPPLSEATTAMGYHLKTLTNFTISNLQYSPDMKGSATFNSTEGVLOHLR 373
DB 459 TPKPATTPLPPLSEATTAMGYHLKTLTNFTISNLQYSPDMKGSATFNSTEGVLOHLR 518
QY 374 PLFOKSSMGPFYLGCOLISLRPEKDAATGVDTTCYHPDPVGPGLDIOQLYWELSQVTH 433
DB 519 PLFOKSSMGPFYLGCOLISLRPEKDAATGVDTTCYHPDPVGPGLDIOQLYWELSQVTH 578
QY 434 GVTOLGFYVLDRLSLFINGYAPONTLSIRGEYQINFHIVMNLSPDPTSSSEYITLLRDIO 493
DB 579 GVTOLGFYVLDRLSLFINGYAPONTLSIRGEYQINFHIVMNLSPDPTSSSEYITLLRDIO 638
QY 494 DKVTLLYKGSQVLDHTRFCLVNTLTMDSVLYTVKALFSSNLDPSLVEQVFLDKTLNASH 553
DB 639 DKVTLLYKGSQVLDHTRFCLVNTLTMDSVLYTVKALFSSNLDPSLVEQVFLDKTLNASH 698
QY 554 WLGSTYQVLDIHYTEMESVYQPTSSSTQHFYLNFTITNLPSYODKAPGTTNORNR 613
DB 699 WLGSTYQVLDIHYTEMESVYQPTSSSTQHFYLNFTITNLPSYODKAPGTTNORNR 758
QY 614 NIEDALNOLFRNSSISYSDCOVSTFRSVPNRHHTGVDSLGNFSLARVDRVAYIEEF 673
DB 759 NIEDALNOLFRNSSISYSDCOVSTFRSVPNRHHTGVDSLGNFSLARVDRVAYIEEF 818
QY 674 LRMTNRGTOLQNTFLDRSSVLYDGYSPNRNEPLTGN 710
DB 819 LRMTNRGTOLQNTFLDRSSVLYDGYSPNRNEPLTGN 855

RESULT 11
US-10-198-053-595
/ Sequence 595, Application US/10198053
/ Patent No. 6858710
/ GENERAL INFORMATION:
/ APPLICANT: Bangur, Chaitanya S.
/ APPLICANT: Recter, Marc W.
/ APPLICANT: Recter, Gary R.
/ APPLICANT: Hall, Paul
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
/ TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
/ FILE REFERENCE: 210121.462C9
/ CURRENT APPLICATION NUMBER: US/10/198,053
/ CURRENT FILING DATE: 2002-07-17
/ NUMBER OF SEQ ID NOS: 624
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 595
/ LENGTH: 3451
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: VARIANT
/ LOCATION: 177, 335, 523, 618, 663, 875, 961, 1001, 1441, 1555, 1560,
/ LOCATION: 1563, 1574, 1585, 2065, 2070, 2683, 2990, 3269, 3381, 3401
/ OTHER INFORMATION: Xaa = Any Amino Acid
US-10-198-053-595

Query Match 92.5%; Score 3651; DB 2; Length 3451;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 694; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 14 FTHRSSVSTSTGPTPTVYIGASKTPASIFGSPASASHLLILFTLNFTITNLRYEENMPG 73
DB 2696 FTHRSSVSTSTGPTPTVYIGASKTPASIFGSPASASHLLILFTLNFTITNLRYEENMPG 2755

QY 74 SRKNTTERVLQGLRLPFXKTSVGPLYSGCRLLTLRPEKDGATGVDALCTHRPDPGP 133
DB 2756 SRKNTTERVLQGLRLPFXKTSVGPLYSGCRLLTLRPEKDGATGVDALCTHRPDPGP 2815

QY 134 GLDREQLYLELSOLTHSITELGPTLDRDSL YNNGFTHRSSVPTSTGVSEEPFTLNFT 193
DB 2816 GLDREQLYLELSOLTHSITELGPTLDRDSL YNNGFTHRSSVPTSTGVSEEPFTLNFT 2875

QY 194 INNLRYADMGQPGSLKFNITDVMKHLSPFORSSLGARYTCRVIALRSYKNGAETR 253
DB 2876 INNLRYADMGQPGSLKFNITDVMKHLSPFORSSLGARYTCRVIALRSYKNGAETR 2935

QY 254 VDLICTYLOPLSGPGLPIKQVFHLSQQTGTRIGPYSIDKDSL YNNGYNEBGPDPPT 313
DB 2936 VDLICTYLOPLSGPGLPIKQVFHLSQQTGTRIGPYSIDKDSL YNNGYNEBGPDPPT 2995

QY 314 TPKPATTFLPPLSEATTAMGYHLKTLTNFTISNLQYSPDMGKSATFNSTEGVLQHLR 373
DB 2996 TPKPATTFLPPLSEATTAMGYHLKTLTNFTISNLQYSPDMGKSATFNSTEGVLQHLR 3055

QY 374 PLFOKSMGPFYIGCOLISLRPEKGAATGVTCTYHPDPVAGGLDIQOLYELSQLTH 433
DB 3056 PLFOKSMGPFYIGCOLISLRPEKGAATGVTCTYHPDPVAGGLDIQOLYELSQLTH 3115

QY 434 GVTQLGFFVYLDRLSLFINGYAPQNL SIRGEYQINFHIVNNLSPNPTSEYITLLADIQ 493
DB 3116 GVTQLGFFVYLDRLSLFINGYAPQNL SIRGEYQINFHIVNNLSPNPTSEYITLLADIQ 3175

QY 494 DKVTLTKGSQLHDTFRFCLVTNL TMDSVLVYKALFSSNLDPDLVEQVFLDKTLNLSFH 553
DB 3176 DKVTLTKGSQLHDTFRFCLVTNL TMDSVLVYKALFSSNLDPDLVEQVFLDKTLNLSFH 3235

QY 554 WLGSTYQVLDIHTEMESYVQPTSSSTQHFYLNFTITNL PYSDKAQPGTNYQNRK 613
DB 3236 WLGSTYQVLDIHTEMESYVQPTSSSTQHFYLNFTITNL PYSDKAQPGTNYQNRK 3255

QY 614 NIEDALNQLFRNSSIKSYFSDQVSTRSVPNRHHTGVDSL CNFSPLARVDRAVAIYEEF 673
DB 3296 NIEDALNQLFRNSSIKSYFSDQVSTRSVPNRHHTGVDSL CNFSPLARVDRAVAIYEEF 3355

QY 674 LRMTNGTQLONFLDRSSVYLDGYSFNRNEPPLTGN 710
DB 3356 LRMTNGTQLONFLDRSSVYLDGYSFNRNEPPLTGN 3392

RESULT 12
US-10-198-053-458
; Sequence 458, Application US/10198053
; Patent No. 6858710
; GENERAL INFORMATION:
; APPLICANT: Bangur, Chaityanya S.
; APPLICANT: Retter, Marc W.
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.462C9
; CURRENT APPLICATION NUMBER: US/10/198.053
; NUMBER OF SEQ ID NOS: 624
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 458
; LENGTH: 1148
; TYPE: PRF
; ORGANISM: Homo sapiens

US-10-198-053-458

Query Match 92.3%; Score 3642; DB 2; Length 1148;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 694; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 14 FTHRSSVSTSTGPTPTVYIGASKTPASIFGSPASASHLLILFTLNFTITNLRYEENMPG 73
DB 393 FTHRSSVSTSTGPTPTVYIGASKTPASIFGSPASASHLLILFTLNFTITNLRYEENMPG 452

QY 74 SRKNTTERVLQGLRLPFXKTSVGPLYSGCRLLTLRPEKDGATGVDALCTHRPDPGP 133
DB 453 SRKNTTERVLQGLRLPFXKTSVGPLYSGCRLLTLRPEKDGATGVDALCTHRPDPGP 512

QY 134 GLDREQLYLELSOLTHSITELGPTLDRDSL YNNGFTHRSSVPTSTGVSEEPFTLNFT 193
DB 513 GLDREQLYLELSOLTHSITELGPTLDRDSL YNNGFTHRSSVPTSTGVSEEPFTLNFT 572

QY 194 INNLRYADMGQPGSLKFNITDVMKHLSPFORSSLGARYTCRVIALRSYKNGAETR 253
DB 573 INNLRYADMGQPGSLKFNITDVMKHLSPFORSSLGARYTCRVIALRSYKNGAETR 632

QY 254 VDLICTYLOPLSGPGLPIKQVFHLSQQTGTRIGPYSIDKDSL YNNGYNEBGPDPPT 313
DB 633 VDLICTYLOPLSGPGLPIKQVFHLSQQTGTRIGPYSIDKDSL YNNGYNEBGPDPPT 692

QY 314 TPKPATTFLPPLSEATTAMGYHLKTLTNFTISNLQYSPDMGKSATFNSTEGVLQHLR 373
DB 693 TPKPATTFLPPLSEATTAMGYHLKTLTNFTISNLQYSPDMGKSATFNSTEGVLQHLR 752

QY 374 PLFOKSMGPFYIGCOLISLRPEKGAATGVTCTYHPDPVAGGLDIQOLYELSQLTH 433
DB 753 PLFOKSMGPFYIGCOLISLRPEKGAATGVTCTYHPDPVAGGLDIQOLYELSQLTH 812

QY 434 GVTQLGFFVYLDRLSLFINGYAPQNL SIRGEYQINFHIVNNLSPNPTSEYITLLADIQ 493
DB 813 GVTQLGFFVYLDRLSLFINGYAPQNL SIRGEYQINFHIVNNLSPNPTSEYITLLADIQ 872

QY 494 DKVTLTKGSQLHDTFRFCLVTNL TMDSVLVYKALFSSNLDPDLVEQVFLDKTLNLSFH 553
DB 873 DKVTLTKGSQLHDTFRFCLVTNL TMDSVLVYKALFSSNLDPDLVEQVFLDKTLNLSFH 932

QY 554 WLGSTYQVLDIHTEMESYVQPTSSSTQHFYLNFTITNL PYSDKAQPGTNYQNRK 613
DB 933 WLGSTYQVLDIHTEMESYVQPTSSSTQHFYLNFTITNL PYSDKAQPGTNYQNRK 992

QY 614 NIEDALNQLFRNSSIKSYFSDQVSTRSVPNRHHTGVDSL CNFSPLARVDRAVAIYEEF 673
DB 993 NIEDALNQLFRNSSIKSYFSDQVSTRSVPNRHHTGVDSL CNFSPLARVDRAVAIYEEF 1052

QY 674 LRMTNGTQLONFLDRSSVYLDGYSFNRNEPPLTGN 710
DB 1053 LRMTNGTQLONFLDRSSVYLDGYSFNRNEPPLTGN 1089

RESULT 13
US-10-198-053-479
; Sequence 479, Application US/10198053
; Patent No. 6858710
; GENERAL INFORMATION:
; APPLICANT: Bangur, Chaityanya S.
; APPLICANT: Retter, Marc W.
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.462C9
; CURRENT APPLICATION NUMBER: US/10/198.053
; NUMBER OF SEQ ID NOS: 624
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 479
; LENGTH: 1148
; TYPE: PRF
; ORGANISM: Homo sapiens

TYPE: PR1
ORGANISM: Homo sapiens
US-10-198-053-479

Query Match 92.3%; Score 3642; DB 2; Length 1148;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 694; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

QY 14 FTHRSVSTSTSPGPTVYLGASKT PASIFGSAASHLLIFLTNFTTNLYEENMMPG 73
DB 393 FTHRSVSTSTSPGPTVYLGASKT PASIFGSAASHLLIFLTNFTTNLYEENMMPG 452
QY 74 SRKFNTTERVLOGLRPLFKNTSVGPLYSGCRLLTLRPEKDEAGVDAICTHRDPDGP 133
DB 453 SRKFNTTERVLOGLRPLFKNTSVGPLYSGSRLLTLRPEKDEAGVDAICTHRDPDGP 512
QY 134 GUDREQLYLELSQTHSITELGPYTLDRDSL YVNGFTHRSVPTSTGVVSEEPFLNPT 193
DB 513 GUDREQLYLELSQTHSITELGPYTLDRDSL YVNGFTHRSVPTSTGVVSEEPFLNPT 572
QY 194 INNLRYMADMGOQGS LKFNITDNVMKHLSP LFORSIGARTGCRVIALRSVKNCAETR 253
DB 573 INNLRYMADMGOQGS LKFNITDNVMKHLSP LFORSIGARTGCRVIALRSVKNCAETR 632
QY 254 VDLCTYLOPLSGPGLPIKOVFHELSQOHTGTRIGPYSLDKDSL YLNGYNEPDPDEPT 313
DB 633 VDLCTYLOPLSGPGLPIKOVFHELSQOHTGTRIGPYSLDKDSL YLNGYNEPDPDEPT 692
QY 314 TPKPATTFPLPSEATTAMGYHLKTLTNFTISNLQYSPDMKGSA TFNSTEGVLQHLR 373
DB 693 TPKPATTFPLPSEATTAMGYHLKTLTNFTISNLQYSPDMKGSA TFNSTEGVLQHLR 752
QY 374 PLFOKSMGPFFYGCQGLSLRPEKGAATGVDTTCYHDPVPGGLDIQOLYELWSQLTH 433
DB 753 PLFOKSMGPFFYGCQGLSLRPEKGAATGVDTTCYHDPVPGGLDIQOLYELWSQLTH 812
QY 434 GVTOLGFVYLDRLDSLFINGYAPONLSIRGEYQINFHIVMNLSPDPTSEYITLLRDIO 493
DB 813 GVTOLGFVYLDRLDSLFINGYAPONLSIRGEYQINFHIVMNLSPDPTSEYITLLRDIO 872
QY 494 DKYTTLYKGSQ LHDTRFCLVTNLTMDSVLYVYKALFSSNLDPDLVEQVFLDKTLNASH 553
DB 873 DKYTTLYKGSQ LHDTRFCLVTNLTMDSVLYVYKALFSSNLDPDLVEQVFLDKTLNASH 932
QY 554 WLGSTYQVLDIHTHESSVYOPTSSSTOHFLNFTTNLPYSQDAQPGTTNYORNR 613
DB 933 WLGSTYQVLDIHTHESSVYOPTSSSTOHFLNFTTNLPYSQDAQPGTTNYORNR 992
QY 614 NIEDALNQLFRNSSISYFSDCOVSTFRSVPNRHHTGVDSL CNFSPILARRVDRVALYEEF 673
DB 993 NIEDALNQLFRNSSISYFSDCOVSTFRSVPNRHHTGVDSL CNFSPILARRVDRVALYEEF 1052
QY 674 LRMTNRGTOLQNFLLDRSSVLYDGYSPNREPLTGN 710
DB 1053 LRMTNRGTOLQNFLLDRSSVLYDGYSPNREPLTGN 1089

```

RESULT 14
US-09-827-271-458
Sequence 458, Application US/09827271
Patent No. 6962980
GENERAL INFORMATION:
APPLICANT: Retter, Marc W.
APPLICANT: Fanger, Gary R.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF OVARIAN CANCER
FILE REFERENCE: 210121.462C6
CURRENT APPLICATION NUMBER: US/09/827.271
CURRENT FILING DATE: 2001-04-04
NUMBER OF SEQ ID NOS: 461
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 458
LENGTH: 1148

TYPE: PR1
ORGANISM: Homo sapiens
US-09-827-271-458

Query Match 92.3%; Score 3642; DB 2; Length 1148;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 694; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

QY 14 FTHRSVSTSTSPGPTVYLGASKT PASIFGSAASHLLIFLTNFTTNLYEENMMPG 73
DB 393 FTHRSVSTSTSPGPTVYLGASKT PASIFGSAASHLLIFLTNFTTNLYEENMMPG 452
QY 74 SRKFNTTERVLOGLRPLFKNTSVGPLYSGCRLLTLRPEKDEAGVDAICTHRDPDGP 133
DB 453 SRKFNTTERVLOGLRPLFKNTSVGPLYSGSRLLTLRPEKDEAGVDAICTHRDPDGP 512
QY 134 GUDREQLYLELSQTHSITELGPYTLDRDSL YVNGFTHRSVPTSTGVVSEEPFLNPT 193
DB 513 GUDREQLYLELSQTHSITELGPYTLDRDSL YVNGFTHRSVPTSTGVVSEEPFLNPT 572
QY 194 INNLRYMADMGOQGS LKFNITDNVMKHLSP LFORSIGARTGCRVIALRSVKNCAETR 253
DB 573 INNLRYMADMGOQGS LKFNITDNVMKHLSP LFORSIGARTGCRVIALRSVKNCAETR 632
QY 254 VDLCTYLOPLSGPGLPIKOVFHELSQOHTGTRIGPYSLDKDSL YLNGYNEPDPDEPT 313
DB 633 VDLCTYLOPLSGPGLPIKOVFHELSQOHTGTRIGPYSLDKDSL YLNGYNEPDPDEPT 692
QY 314 TPKPATTFPLPSEATTAMGYHLKTLTNFTISNLQYSPDMKGSA TFNSTEGVLQHLR 373
DB 693 TPKPATTFPLPSEATTAMGYHLKTLTNFTISNLQYSPDMKGSA TFNSTEGVLQHLR 752
QY 374 PLFOKSMGPFFYGCQGLSLRPEKGAATGVDTTCYHDPVPGGLDIQOLYELWSQLTH 433
DB 753 PLFOKSMGPFFYGCQGLSLRPEKGAATGVDTTCYHDPVPGGLDIQOLYELWSQLTH 812
QY 434 GVTOLGFVYLDRLDSLFINGYAPONLSIRGEYQINFHIVMNLSPDPTSEYITLLRDIO 493
DB 813 GVTOLGFVYLDRLDSLFINGYAPONLSIRGEYQINFHIVMNLSPDPTSEYITLLRDIO 872
QY 494 DKYTTLYKGSQ LHDTRFCLVTNLTMDSVLYVYKALFSSNLDPDLVEQVFLDKTLNASH 553
DB 873 DKYTTLYKGSQ LHDTRFCLVTNLTMDSVLYVYKALFSSNLDPDLVEQVFLDKTLNASH 932
QY 554 WLGSTYQVLDIHTHESSVYOPTSSSTOHFLNFTTNLPYSQDAQPGTTNYORNR 613
DB 933 WLGSTYQVLDIHTHESSVYOPTSSSTOHFLNFTTNLPYSQDAQPGTTNYORNR 992
QY 614 NIEDALNQLFRNSSISYFSDCOVSTFRSVPNRHHTGVDSL CNFSPILARRVDRVALYEEF 673
DB 993 NIEDALNQLFRNSSISYFSDCOVSTFRSVPNRHHTGVDSL CNFSPILARRVDRVALYEEF 1052
QY 674 LRMTNRGTOLQNFLLDRSSVLYDGYSPNREPLTGN 710
DB 1053 LRMTNRGTOLQNFLLDRSSVLYDGYSPNREPLTGN 1089

```

RESULT 15
US-10-198-053-459
Sequence 459, Application US/10198053
Patent No. 6858710
GENERAL INFORMATION:
APPLICANT: Bangur, Chaitanya S.
APPLICANT: Retter, Marc W.
APPLICANT: Fanger, Gary R.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
FILE REFERENCE: 210121.462C9
CURRENT APPLICATION NUMBER: US/10/198.053
CURRENT FILING DATE: 2002-07-17
NUMBER OF SEQ ID NOS: 624
SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 459
; LENGTH: 1156
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-198-053-459

Query Match 92.3%; Score 1642; DB 2; length 1156;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 694; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

| | | | |
|----|------|---|------|
| QY | 14 | FTHRSSVSTSTPTGPTVYLGAASKTPASIFGPSASHLLIFTLNFTITNLRYEENMPG | 73 |
| DB | 401 | FTHRSSVSTSTPTGPTVYLGAASKTPASIFGPSASHLLIFTLNFTITNLRYEENMPG | 460 |
| QY | 74 | SRKFNTERVLOGLRPLFKNTSVGLYSGCRLTLARPEKDEATGVDAICTHRPDPGP | 133 |
| DB | 461 | SRKFNTERVLOGLRPLFKNTSVGLYSGCRLTLARPEKDEATGVDAICTHRPDPGP | 520 |
| QY | 134 | GLDBEQLYLELSQLTSHITELGPTLDRDSLTVNGFTHRSSVPTSTGWSSEPPFTLNT | 193 |
| DB | 521 | GLDBEQLYLELSQLTSHITELGPTLDRDSLTVNGFTHRSSVPTSTGWSSEPPFTLNT | 580 |
| QY | 194 | INNLRYADMGOPGSLKFNITDNVMKHLSPLFQSSIGARYTCRVIALRSVNGAETR | 253 |
| DB | 581 | INNLRYADMGOPGSLKFNITDNVMKHLSPLFQSSIGARYTCRVIALRSVNGAETR | 640 |
| QY | 254 | VDLCTYLQPLSGGGLPIKOVFHELSQQTGITRLGYSIDKDSLTVNGYNEPGLDEPPT | 313 |
| DB | 641 | VDLCTYLQPLSGGGLPIKOVFHELSQQTGITRLGYSIDKDSLTVNGYNEPGLDEPPT | 700 |
| QY | 314 | TPKPATTEFLPLSATATAMGVHLKTLTLNFTISNLQYSPDMGKSATFNSTEGVLOHLR | 373 |
| DB | 701 | TPKPATTEFLPLSATATAMGVHLKTLTLNFTISNLQYSPDMGKSATFNSTEGVLOHLR | 760 |
| QY | 374 | PLFQKSMGPFYLGCCQLISLRPEKGAATGVDITCTYHPDPVGGGLDIQOLYWELSQLT | 433 |
| DB | 761 | PLFQKSMGPFYLGCCQLISLRPEKGAATGVDITCTYHPDPVGGGLDIQOLYWELSQLT | 820 |
| QY | 434 | GVTQLGFYVLDRLSLFTNGYAPQNLIRGEYQINFHIVNNLSNPDPTSSEYITLLRDIQ | 493 |
| DB | 821 | GVTQLGFYVLDRLSLFTNGYAPQNLIRGEYQINFHIVNNLSNPDPTSSEYITLLRDIQ | 880 |
| QY | 494 | DKYTTLYKGSQQLDTPFECVLTNLTMDSVLVYKALFSSNLDPSLYEQVFLDKTLNASFH | 553 |
| DB | 881 | DKYTTLYKGSQQLDTPFECVLTNLTMDSVLVYKALFSSNLDPSLYEQVFLDKTLNASFH | 940 |
| QY | 554 | WLGSYQVLDIHTEMESSYQPTSSSTQHFYLNFTITNLPSQDKAOPGTINYQNRK | 613 |
| DB | 941 | WLGSYQVLDIHTEMESSYQPTSSSTQHFYLNFTITNLPSQDKAOPGTINYQNRK | 1000 |
| QY | 614 | NIEDALNQLFRNSSIKSYFSDCQVSTFRSVPNRHHTGVDSL CNFSPLARRVDRVAIYEEF | 673 |
| DB | 1001 | NIEDALNQLFRNSSIKSYFSDCQVSTFRSVPNRHHTGVDSL CNFSPLARRVDRVAIYEEF | 1060 |
| QY | 674 | LRMTTRNGTQLONTFLDRSSVLVDGYSPNRNEPLTGNS | 710 |
| DB | 1061 | LRMTTRNGTQLONTFLDRSSVLVDGYSPNRNEPLTGNS | 1097 |

Search completed: October 14, 2006, 04:06:07
Job time : 57 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: October 14, 2006, 03:55:05 ; Search time 202 Seconds
(without alignments)
1693.059 Million cell updates/sec

Title: US-10-687-035-1
Perfect score: 3945
Sequence: 1 AAGPARARARTKLFTTRSSV.....QKLISEDLMTGHHHHH 748

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2589679 seqs, 457216429 residues

Total number of hits satisfying chosen parameters: 2589679

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq.8:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*
10: geneseqp2006s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|----------|--------------------|
| 1 | 3945 | 100.0 | 748 | ADS94302 | AdS94302 CA 125/O7 |
| 2 | 3904.5 | 99.0 | 809 | ADS94303 | AdS94303 CA 125/O7 |
| 3 | 3665 | 92.9 | 6995 | AEC60057 | Aec60057 Human muc |
| 4 | 3663 | 92.9 | 833 | AA812554 | Aa812554 Human ova |
| 5 | 3663 | 92.9 | 833 | ABP30896 | Abp30896 0772P c10 |
| 6 | 3663 | 92.9 | 833 | ADA08542 | Ada08542 Human ova |
| 7 | 3663 | 92.9 | 833 | ADP08885 | Adp08885 Secreted |
| 8 | 3663 | 92.9 | 833 | ADG46172 | Adg46172 Human ova |
| 9 | 3663 | 92.9 | 833 | ADX17768 | Adx17768 Human ova |
| 10 | 3663 | 92.9 | 914 | AA812552 | Aa812552 Human ova |
| 11 | 3663 | 92.9 | 914 | AA899203 | Aa899203 Human ova |
| 12 | 3663 | 92.9 | 914 | ABP30968 | Abp30968 Hypotheti |
| 13 | 3663 | 92.9 | 914 | ABP30896 | Abp30896 0772P pro |
| 14 | 3663 | 92.9 | 914 | ADU01425 | Adu01425 Breast ca |
| 15 | 3663 | 92.9 | 914 | ADZ41689 | Adz41689 Human bre |
| 16 | 3663 | 92.9 | 914 | ADA08631 | Ada08631 Human o77 |
| 17 | 3663 | 92.9 | 914 | ADA08465 | Ada08465 Human ova |
| 18 | 3663 | 92.9 | 914 | ADP08974 | Adp08974 Secreted |
| 19 | 3663 | 92.9 | 914 | ADP08808 | Adp08808 Secreted |
| 20 | 3663 | 92.9 | 914 | ADG46095 | Adg46095 Human ova |
| 21 | 3663 | 92.9 | 914 | ADN40451 | Adn40451 Human bre |
| 22 | 3663 | 92.9 | 914 | ADX17691 | Adx17691 Human ova |
| 23 | 3663 | 92.9 | 914 | ADX17857 | Adx17857 Partial h |

| | | | | | | |
|----|--------|------|-------|----|----------|--------------------|
| 24 | 3658 | 92.7 | 1889 | 6 | AAE34700 | Aae34700 Protein e |
| 25 | 3658 | 92.7 | 1890 | 5 | ABG96381 | Abg96381 Human ova |
| 26 | 3658 | 92.7 | 1890 | 8 | ADM12352 | Adm12352 Human CA1 |
| 27 | 3658 | 92.7 | 1890 | 10 | AEE01059 | Aee01059 Ovarian c |
| 28 | 3655.5 | 92.7 | 1890 | 8 | ADO38574 | Ado38574 Ovarian c |
| 29 | 3651 | 92.5 | 3451 | 5 | ABP31026 | Abp31026 Amino aci |
| 30 | 3651 | 92.5 | 3451 | 7 | ADA08748 | Ada08748 Human O77 |
| 31 | 3651 | 92.5 | 3451 | 7 | ADP09091 | Adp09091 Secreted |
| 32 | 3651 | 92.5 | 3451 | 9 | ADX17974 | Adx17974 Human ova |
| 33 | 3649 | 92.5 | 1148 | 6 | ABU54692 | Abu54692 Human CA1 |
| 34 | 3649 | 92.5 | 2248 | 6 | ABU54693 | Abu54693 Human CA1 |
| 35 | 3649 | 92.5 | 11721 | 6 | ABU54721 | Abu54721 Human CA1 |
| 36 | 3649 | 92.5 | 22157 | 8 | ADP84155 | Adp84155 Human CA1 |
| 37 | 3646 | 92.4 | 1148 | 6 | ABU00145 | Abu00145 Human nov |
| 38 | 3642 | 92.3 | 1148 | 4 | ABBS5836 | Abbs5836 Human pro |
| 39 | 3642 | 92.3 | 1148 | 4 | ABBS5823 | Abbs5823 HOST-1 ov |
| 40 | 3642 | 92.3 | 1148 | 5 | ABG96380 | Abg96380 Human ova |
| 41 | 3642 | 92.3 | 1148 | 5 | ABP30964 | Abp30964 Truncated |
| 42 | 3642 | 92.3 | 1148 | 5 | ABP30969 | Abp30969 Clone FLJ |
| 43 | 3642 | 92.3 | 1148 | 7 | ADA08632 | Ada08632 Human O77 |
| 44 | 3642 | 92.3 | 1148 | 7 | ADA08611 | Ada08611 Human O77 |
| 45 | 3642 | 92.3 | 1148 | 7 | ADP08954 | Adp08954 Secreted |

ALIGNMENTS

RESULT 1
ID ADS94302 standard; protein; 748 AA.
XX
AC ADS94302:
XX
XX
XX 02-DEC-2004 (first entry)
XX
XX CA 125/0772P 3-repeat amino acid sequence SEQ ID NO:1.
DE
XX
XX antibody; antigen-binding antibody fragment;
KW cell-associated CA 125/0772P; monoclonal antibody; cytostatic;
KW immunostimulant; mediator of lysis; tumour; cell proliferative disorder;
KW cancer; cervical cancer; uterine cancer; breast cancer; lung cancer;
KW ovarian cancer.
XX
XX
OS Synthetic.
XX
XX WO2004035537-A2.
XX
XX PN 29-APR-2004.
XX
XX PD
XX
XX PF 15-OCT-2003; 2003WO-US032945.
XX
XX PR 16-OCT-2002; 2002US-0418828P.
XX
XX PR 10-JUL-2003; 2003US-0485986P.
XX
XX PA (EURO-) EUROCELLTQUE SA.
XX
XX PI Albone EF, Soltis DA;
XX
XX WPI; 2004-357171/33.
XX
XX DR Novel isolated antibody, or antigen-binding antibody fragment binding
XX with cell-associated CA 125/0772P polypeptide relative to shed CA
XX PT 125/0772 polypeptide, useful for ameliorating cervical or ovarian cancer.
XX
XX PS Example; SEQ ID NO 1; 153pp; English.
XX
XX CC The present invention describes an isolated antibody, or an antigen-
XX binding antibody fragment (I), that preferentially binds cell-associated
XX CA 125/0772P polypeptide relative to shed CA 125/0772 polypeptide. Also
XX described: (1) a monoclonal antibody (II) produced by hybridoma 4E7; (2)
XX a monoclonal antibody that competes with binding of (II); (3) a hybridoma
XX as deposited in (II); (4) an isolated nucleic acid molecule (III)
XX comprising a nucleotide sequence that encodes a variable chain region of

(1); (5) a pharmaceutical composition comprising an antibody or an antigen-binding antibody fragment that preferentially binds cell-associated CA 125/0772p polypeptide relative to shed CA 125/0772p polypeptide, and a carrier; (6) a pharmaceutical composition comprising a monoclonal antibody or an antigen-binding monoclonal antibody fragment that preferentially binds cell-associated CA 125/0772p polypeptide relative to shed CA 125/0772p polypeptide, and a carrier; (7) an article of manufacture (IV) comprising packaging material and a composition comprising an antibody, or an antigen-binding antibody fragment that preferentially binds cell-associated CA 125/0772p polypeptide relative to shed CA 125/0772p polypeptide, and a carrier; (8) a fusion polypeptide (V) comprising an antibody, or an antigen-binding antibody fragment, which preferentially binds cell-associated CA 125/0772p polypeptide relative to shed CA 125/0772p polypeptide, and a carrier; (9) a monoclonal antibody (VI) chosen from 325.1, 621.1, 633.1, 654.1, 725.1, 8G9, 7F10, 8A1, 8AC3, 15C9, 8B5, 7G10, 16C7, 7C6, 7H1, 16H9, 7A11, 4E7, 117.1, 368.1, 446.1, 501.1, and 776.1 or its antigen-binding antibody fragment; (11) an antibody or antigen binding antibody fragment that competes with (VI); and (12) a pharmaceutical composition comprising (VI) and a carrier. (I) has cytostatic activity, and can be used as an immunostimulant and a mediator of lysis of positive tumour cell. (I) is useful for ameliorating a symptom of a CA 125/077p-related disorder which is a cell proliferative disorder such as cancer, cervical or uterine cancer, breast or lung cancer or ovarian cancer. (V) is useful for diagnosing a symptom of a CA 125/0772p polypeptide relative to shed CA 125/0772p polypeptide, and a carrier. (9) is useful for monitoring the development or progression of cancer or tumour as part of clinical testing procedure. The present sequence represents the CA 125/0772p 3-repeat amino acid sequence, which is used in the exemplification of the present invention.

Seq Sequence 748 AA:

Query Match 100.0%; Score 3945; DB 8; Length 748;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 748; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AADPARARATKFTFRSSVSTSTSTGTPTVYGAATPAATGPGSAASHLLFTLNFT 60
DB 1 AADPARARATKFTFRSSVSTSTSTGTPTVYGAATPAATGPGSAASHLLFTLNFT 60
QY 61 ITNLRYBENWPGSRKFNTERVLOGLRPLFNNTSGVPLYSQCRLTLRPEKDGATGV 120
DB 61 ITNLRYBENWPGSRKFNTERVLOGLRPLFNNTSGVPLYSQCRLTLRPEKDGATGV 120
QY 121 DAICTHRPDPGTGDLREQLYLELSQITLSTELGPTLDRSLVYNGFTFRSSVPTTST 180
DB 121 DAICTHRPDPGTGDLREQLYLELSQITLSTELGPTLDRSLVYNGFTFRSSVPTTST 180
QY 181 GVUSEBPTLNTFTNNLRVYADMGQPGSLKFNITDVMKHLSPLRQSSLGARYGCRV 240
DB 181 GVUSEBPTLNTFTNNLRVYADMGQPGSLKFNITDVMKHLSPLRQSSLGARYGCRV 240
QY 241 IARSVNGAEKRVVDLCYLOPLSGPGLPIKQVFEHLSQOCHGIRLGLGYSLDKSLYL 300
DB 241 IARSVNGAEKRVVDLCYLOPLSGPGLPIKQVFEHLSQOCHGIRLGLGYSLDKSLYL 300
QY 301 NGYNEBPDDEPTTPKATFTLPLSEATTAMGVHLKTLTNTFTISNLQYSPDMKGSAT 360
DB 301 NGYNEBPDDEPTTPKATFTLPLSEATTAMGVHLKTLTNTFTISNLQYSPDMKGSAT 360
QY 361 FNSBEGVLQHLRPLFOKSSMGPPYLCQQLISLRPEKDGATGVDTCTHPPVGGGLD 420
DB 361 FNSBEGVLQHLRPLFOKSSMGPPYLCQQLISLRPEKDGATGVDTCTHPPVGGGLD 420
QY 421 IQLYMEISQLTGVQLGFYVLDRLFTNGVAPQWLSTRGELQIFHIVNNLSRDP 480
DB 421 IQLYMEISQLTGVQLGFYVLDRLFTNGVAPQWLSTRGELQIFHIVNNLSRDP 480
QY 481 TSSEYITLADIODKVTTLTKGSQLHDTFFCLVTNLTMSDVLTVAALFSSNLDPSTVE 540
DB 481 TSSEYITLADIODKVTTLTKGSQLHDTFFCLVTNLTMSDVLTVAALFSSNLDPSTVE 540

QY 541 QVFLDKTLNASEFHWLGSSTYQVLDIHVTEMESSVYPTSSSSSTQHFYLNFTINLPSQDK 600
DB 541 QVFLDKTLNASEFHWLGSSTYQVLDIHVTEMESSVYPTSSSSSTQHFYLNFTINLPSQDK 600
QY 601 AQPSTTNTQNRKRNIEDALNQLFNSSISKYSFSDCVSTFRSVPRRHITGVDSLGNFSL 660
DB 601 AQPSTTNTQNRKRNIEDALNQLFNSSISKYSFSDCVSTFRSVPRRHITGVDSLGNFSL 660
QY 661 ARVRVVAIYEFELMTNNGTQLONFPLDRSSVLVDGYSPPNNEPLTGNADIOHSGGRS 720
DB 661 ARVRVVAIYEFELMTNNGTQLONFPLDRSSVLVDGYSPPNNEPLTGNADIOHSGGRS 720
QY 721 SLEGPFQKLSIEDLNMTGHHHHH 748
DB 721 SLEGPFQKLSIEDLNMTGHHHHH 748
RESULT 2
ID ADS94303 standard; protein; 809 AA.
XX ADS94303;
AC ADS94303;
XX 02-DEC-2004 (fixet entry)
XX CA 125/0772p 3-repeat TM amino acid sequence SEQ ID NO:2.
DE antibody; antigen-binding antibody fragment;
XX cell-associated CA 125/0772p; monoclonal antibody; cytostatic;
XX immunostimulant; mediator of lysis; tumour; cell proliferative disorder;
XX cancer; cervical cancer; uterine cancer; breast cancer; lung cancer;
XX ovarian cancer.
XX Synthetic.
OS WO2004035537-A2.
XX MO2004035537-A2.
XX 29-APR-2004.
PD 15-OCT-2003; 2003MO-US032945.
XX 15-OCT-2003; 2003MO-US032945.
PR 16-OCT-2002; 2002US-0418828P.
XX 10-JUL-2003; 2003US-0485986P.
XX (EURO-) EUROCELLTIGUE SA.
PA Albone EF, Solits DA;
PI WPI; 2004-357171/33.
DR Novel isolated antibody, or antigen-binding antibody fragment binding
XX with cell-associated CA 125/0772p polypeptide relative to shed CA
PT 125/0772 polypeptide, useful for ameliorating cervical or ovarian cancer.
XX Example; SEQ ID NO 2; 153pp; English.
PS The present invention describes an isolated antibody, or an antigen-
XX binding antibody fragment (I), that preferentially binds cell-associated
XX CA 125/0772p polypeptide relative to shed CA 125/0772 polypeptide. Also
XX described: (1) a monoclonal antibody (II) produced by hybridoma 4E7; (2)
XX a monoclonal antibody that competes with binding of (II); (3) a hybridoma
XX as deposited in (II); (4) an isolated nucleic acid molecule (III)
XX comprising a nucleotide sequence that encodes a variable chain region of
XX (I); (5) a pharmaceutical composition comprising an antibody or an
XX antigen-binding antibody fragment that preferentially binds cell-
XX associated CA 125/0772p polypeptide relative to shed CA 125/0772p
XX polypeptide, and a carrier; (6) a pharmaceutical composition comprising a
XX monoclonal antibody or an antigen-binding monoclonal antibody fragment
XX that preferentially binds cell-associated CA 125/0772p polypeptide
XX relative to shed CA 125/0772p polypeptide, and a carrier; (7) an article
XX of manufacture (IV) comprising packaging material and a composition
XX comprising an antibody, or an antigen-binding antibody fragment that
XX preferentially binds cell-associated CA 125/0772p relative to shed CA

CC 125/0772P, and a carrier contained within the packaging material, and
 CC composition in a form suitable for administration to a subject; (8) a
 CC fusion polypeptide (V) comprising an antibody, or an antigen-binding
 CC antibody fragment, which preferentially binds cell-associated CA
 CC 125/0772P relative to shed CA 125/0772P operably linked to a heterologous
 CC agent; (9) ameliorating (M1) a symptom of a CA 125/0772P-related disorder
 CC ; (10) a monoclonal antibody (V1) chosen from 325.1, 621.1, 631.1, 654.1,
 CC 725.1, 869, 7F10, 8A1, 8AC3, 15C9, 8E3, 8B5, 7G10, 16C7, 7G6, 7H1, 16H9,
 CC 7A11, 4E7, 117.1, 368.1, 446.1, 501.1, and 776.1 or its antigen-binding
 CC antibody fragment; (11) an antibody or antigen binding antibody fragment
 CC that competes with (V1) and (12) a pharmaceutical composition comprising
 CC (V1) and a carrier. (I) has cytostatic activity, and can be used as an
 CC immunostimulant and a mediator of lysis of positive tumour cell. (I) is
 CC useful for ameliorating a symptom of a CA 125/077P-related disorder which
 CC is a cell proliferative disorder such as cancer, cervical or uterine
 CC cancer, breast or lung cancer or ovarian cancer. (V) is useful
 CC diagnostically for monitoring the development or progression of cancer or
 CC tumour as part of clinical testing procedure. The present sequence
 CC represents the CA 125/0772P 3-repeat TM amino acid sequence, which is
 CC used in the exemplification of the present invention.

XX Sequence 809 AA:

Query Match 99.0%; Score 3904.5; DB 8; Length 809;
 Best Local Similarity 92.5%; Pred. No. 0;
 Matches 746; Conservative 0; Mismatches 0; Indels 61; Gaps 1;

QY 1 AADPARARRTKLFTHRSSVSTSTPGTPVYLGAATPAISFGPSAASHLLIFLTNFT 60
 DB 1 AADPARARRTKLFTHRSSVSTSTPGTPVYLGAATPAISFGPSAASHLLIFLTNFT 60
 QY ITNLRVENMMPSGRKNTTERVVGILRLPKTSVGPVSGGRLTLRPEKGEATGV 120
 DB ITNLRVENMMPSGRKNTTERVVGILRLPKTSVGPVSGGRLTLRPEKGEATGV 120
 QY 121 DALCTHRPDPYGPDLREQLYLELSQTLSTTELGPTLRDLSLYVNGFTHRSSVPTST 180
 DB 121 DALCTHRPDPYGPDLREQLYLELSQTLSTTELGPTLRDLSLYVNGFTHRSSVPTST 180
 QY 181 GVVESEPTLNTNNLRVYADMGQPSGLKFNITDNMKHLSPLFORSSIGARTCYRV 240
 DB 181 GVVESEPTLNTNNLRVYADMGQPSGLKFNITDNMKHLSPLFORSSIGARTCYRV 240
 QY 241 IALRSVNGAETRDLLCTYLQPLSGPLPKQVFNHLSQCTHGITLGPYSLDKSLYL 300
 DB 241 IALRSVNGAETRDLLCTYLQPLSGPLPKQVFNHLSQCTHGITLGPYSLDKSLYL 300
 QY 301 NGYNEPDEPPTTPKPAITFLPRLSEATAMGVHKLTLNFTISMLQYSPDMKGSAT 360
 DB 301 NGYNEPDEPPTTPKPAITFLPRLSEATAMGVHKLTLNFTISMLQYSPDMKGSAT 360
 QY 361 FNSTEGVLQHLRLPLFKSSMGPFYLCOLISLRPEKOGATGVDTCTYHPPVFGGLD 420
 DB 361 FNSTEGVLQHLRLPLFKSSMGPFYLCOLISLRPEKOGATGVDTCTYHPPVFGGLD 420
 QY 421 IQOYLWELSQLTHGVTOLGFTVLDRLSFLINGTAPQMLSTRGEYQINFIHVNMLSNPD 480
 DB 421 IQOYLWELSQLTHGVTOLGFTVLDRLSFLINGTAPQMLSTRGEYQINFIHVNMLSNPD 480
 QY 481 TSSSEYTLRLDIDQKVTLLKSGQLHDTFFCLVTNLTMSVLTVAALFSSNLDPELVE 540
 DB 481 TSSSEYTLRLDIDQKVTLLKSGQLHDTFFCLVTNLTMSVLTVAALFSSNLDPELVE 540
 QY 541 QVEFLDKTLNLSFHWLGSSTYQVLDIHVTEMSSVYQPTSSSTQHFYLNFTITNLPSYQDK 600
 DB 541 QVEFLDKTLNLSFHWLGSSTYQVLDIHVTEMSSVYQPTSSSTQHFYLNFTITNLPSYQDK 600
 QY 601 AAGCTTYYQNRKNIBALNQLFNSSIKSYFSCQVSTRSVYNNRHHTGVDSICNFSPL 660
 DB 601 AAGCTTYYQNRKNIBALNQLFNSSIKSYFSCQVSTRSVYNNRHHTGVDSICNFSPL 660
 QY 661 AARDRVAIYEEFLMRNNGTOLQNFLLDRSSVYVDGYSNNRNEPLNG----- 708
 DB 661 AARDRVAIYEEFLMRNNGTOLQNFLLDRSSVYVDGYSNNRNEPLNG----- 708

DB 661 AARDRVAIYEEFLMRNNGTOLQNFLLDRSSVYVDGYSNNRNEPLNGSDLPFMAVILI 720
 QY 709 -----NSADIQHSGR 719
 DB 721 GLAGLLGLITCLICGVLVTRRRKKEGEYVQQCPGYOSHLDLEDLQNSADIQHSGR 780
 QY 720 SSLEGPFEQKLISEEDLNAMHTGHHHH 748
 DB 781 SSLEGPFEQKLISEEDLNAMHTGHHHH 809

RESULT 3
 AEC60057
 ID AEC60057 standard; protein; 6995 AA.
 XX AC AEC60057;
 XX DT 17-NOV-2005 (first entry)
 XX DE Human mucin 16 (MUC16).
 XX KW Receptor; tumor marker; drug delivery; drug screening; cancer;
 KW autoimmune disease; infection; Cytostatic; Antimicrobial;
 KW immunosuppressive; Antibacterial; Virocidic; Anti-HIV; Fungicide;
 KW Antiparasitic; Procoagulant; Antimalarial; Anabolic; Hypertensive;
 KW Antiarthritic; Antiallergic; Antiinflammatory; Ophthalmological;
 KW Urothelial; Dermatological.
 XX OS Homo sapiens.
 XX PN WC02005081711-A2.
 XX PD 09-SEP-2005.
 XX PF 05-NOV-2004; 2004MO-US038392.
 XX PR 06-NOV-2003; 2003US-0518534P.
 PR 26-MAR-2004; 2004US-0557116P.
 PR 04-AUG-2004; 2004US-0598899P.
 PR 27-OCT-2004; 2004US-0622455P.
 XX PA (SEAT-) SEATTLE GENETICS INC.
 XX PI Dronina SO, Senter PD, Toki BE, Edens AJ, Kline TB, Polakis P;
 PI Sliwkowski MX, Spencer SD;
 XX DR WPI; 2005-658310/67.
 XX DR GENBANK; AAK74120.
 PT Novel monomethylvaline conjugate compound or its salt or solvate, useful
 PT for treating cancer, autoimmune disease and infectious disease and for
 PT diagnosing cancer.
 XX PS Disclosure; SEQ ID NO 4; 426pp; English.
 XX CC The invention relates to a monomethylvaline conjugate compound (C1) or
 CC its salt or solvate, comprising a ligand unit (particularly an antibody
 CC directed against a disease antigen) and a drug moiety (of formulae
 CC detailed in the specification). Also included are an antibody-drug
 CC conjugate compound (DC) (comprising an antibody covalently attached to
 CC one or more drug moieties), a pharmaceutical composition (PC) comprising
 CC C1 or DC, inhibiting (M1) the growth of tumor cells that overexpress a
 CC tumor-associated antigen (involves administering DC and a
 CC chemotherapeutic agent, to the patient, where DC and the chemotherapeutic
 CC agent are each administered in amounts effective to inhibit growth of
 CC tumor cells in the patient), treating a human patient susceptible to or
 CC diagnosed with a disorder characterized by overexpression of ErbB2
 CC receptor (involving administering a combination of DC and a
 CC chemotherapeutic agent), an article (AM) of manufacture comprising C1 or
 CC DC (comprising a container and a package insert or label indicating that
 CC the C1 or DC can be used to treat cancer characterized by the
 CC overexpression of at least one of CD30, CD40, CD70 and Lewis Y, or to
 CC treat cancer characterized by overexpression of an ErbB2 receptor) and

(M2) cancer involves administering to a patient a formulation of DC and a diluent, carrier or excipient. In C1 or DC, a substantial amount of the drug moiety is not cleaved from the antibody until C1 or DC enters a cell with a cell-surface receptor specific for the antibody of C1 or DC, and the drug moiety is cleaved from the antibody when C1 or DC does enter the cell. The bioavailability of the compound or an intracellular metabolite of the compound in a mammal is improved when compared to a drug compound comprising the drug moiety of C1 or DC. The bioavailability of the compound or an intracellular metabolite of the compound in a mammal is improved when compared to an analog of the compound not having the drug moiety. The drug moiety is intracellularly cleaved in a mammal from the antibody of C1 or DC, or an intracellular metabolite of C1 or DC. The antibody is at least one of AC10, 52C6, BR96, 1F6 and 2F2. In DC, the antibody specifically binds to a HER2 receptor. DC specifically binds to the extracellular domain of the HER2 receptor. DC inhibits growth of tumor cells, which overexpress HER2 receptor. The drug conjugates are useful for killing or inhibiting the proliferation of tumor cells or cancer cells, which involves treating tumor cells or cancer cells with CP or DC, or its salt or solvate. The cancer is characterized by the overexpression of an ErbB2 receptor at a +2 level or above. The drug conjugates are useful for treating an autoimmune disease or infectious disease, (e.g. bacterial infections caused by microorganisms such as Streptococcus pyogenes and Clostridium botulinum, viral infections caused by viruses e.g. Herpes simplex virus and HIV, fungal diseases (e.g. Aspergilliosis and Histoplasmosis), parasitic diseases (e.g. malaria and Chagas disease), and autoimmune diseases such as Addison's disease, arthritis, allergy, Crohn's disease, Reiter's syndrome, and systemic lupus erythematosus), all chosen from an extensive list given in the specification. The present sequence is a human tumor-associated receptor protein, antibodies against which are used in the drug conjugates of the invention.

Sequence 6995 AA;

| | | | | |
|-----------------------|--------------|---------------------|---------------|--------------|
| Query Match | 92.9% | Score 3665; | DB 9; | Length 6995; |
| Best Local Similarity | 99.9% | Pred. NO. 5.2e-314; | | |
| Matches 696; | Conservative | 1; | Mismatches 0; | Indels 0; |
| | | | Gaps | 0 |

| | | | | | |
|----|------|--------------------------------|------------------------------------|--------------|------|
| Qy | 14 | FPHRSVSSTSPGPPYVLAASKTPAI | PPPSAASHLLIFPTNFI | ITNLRYEANMPG | 73 |
| Db | 6240 | FPHRSVSSTSPGPPYVLAASKTPAI | PPPSAASHLLIFPTNFI | ITNLRYEANMPG | 6299 |
| Qy | 74 | SKKFTTERVLOGLIRPLPKNTSVGLYSGCR | TLTRPKGEGATGDAICTHAPDPTGP | 133 | |
| Db | 6300 | SKKFTTERVLOGLIRPLPKNTSVGLYSGCR | TLTRPKGEGATGDAICTHAPDPTGP | 6359 | |
| Qy | 134 | GJDRQOLVIELSOLHSTITELGPYLDLDS | LYVNGFTHRSVPTSTGVSEEPFTNFI | 193 | |
| Db | 6360 | GJDRQOLVIELSOLHSTITELGPYLDLDS | LYVNGFTHRSVPTSTGVSEEPFTNFI | 6419 | |
| Qy | 194 | INNLRYMDMGPGSLKNTITDNVKKHLLSP | FORSSIGANVTGCRVIALRSVNGAETR | 253 | |
| Db | 6420 | INNLRYMDMGPGSLKNTITDNVKKHLLSP | FORSSIGANVTGCRVIALRSVNGAETR | 6479 | |
| Qy | 254 | VLLCTCYLOPLSGPGLPKOVFHEHSOOTH | GITRLGPVLSLKDLSLYNNGNVEPGDDEPT | 313 | |
| Db | 6480 | VLLCTCYLOPLSGPGLPKOVFHEHSOOTH | GITRLGPVLSLKDLSLYNNGNVEPGDDEPT | 6539 | |
| Qy | 314 | TEPKATTEPLPSEATATWAGYHAKTLT | INFTISNLQVSPDMGKSATFNSTEGVLOHLR | 373 | |
| Db | 6540 | TEPKATTEPLPSEATATWAGYHAKTLT | INFTISNLQVSPDMGKSATFNSTEGVLOHLR | 6599 | |
| Qy | 374 | PLPQKSSMGPFLGCOLISLRPEKOGAAG | CVUTTCYHPPDPGRLDIOQLYWEHSOLTH | 433 | |
| Db | 6600 | PLPQKSSMGPFLGCOLISLRPEKOGAAG | CVUTTCYHPPDPGRLDIOQLYWEHSOLTH | 6659 | |
| Qy | 434 | GVTQOLGFVYLDLDSLFLINGYAPONT | SIRGEYOINFIHVMNLSNDPSTSEBITLARDIQ | 493 | |
| Db | 6660 | GVTQOLGFVYLDLDSLFLINGYAPONT | SIRGEYOINFIHVMNLSNDPSTSEBITLARDIQ | 6719 | |
| Qy | 494 | DKYTTILYKKSQJHDFRFLCYTNLTMD | SVLYVYKALLFSFNLDPSLVEOFLDKTLTNAEFH | 553 | |
| Db | 6720 | DKYTTILYKKSQJHDFRFLCYTNLTMD | SVLYVYKALLFSFNLDPSLVEOFLDKTLTNAEFH | 6779 | |

| | | | |
|----|------|--|------|
| Qy | 554 | WLSTYQLONDIHVHEMSSVYQPTSSSSTOHFNFTITLN.PYQODKQOPGTNYQQRKR | 613 |
| Db | 6780 | WLSTYQLOVDIHVHEMSSVYQPTSSSSTOHFNFTITLN.PYQODKQOPGTNYQQRKR | 683 |
| Qy | 614 | NIEBALQOLFNNSSIKSYFSDCOVSTFSPVNRHHTGVDSL.CNFSPLARVDRAVYEEF | 673 |
| Db | 6640 | NIEBALQOLFNNSSIKSYFSDCOVSTFSPVNRHHTGVDSL.CNFSPLARVDRAVYEEF | 689 |
| Qy | 674 | LRMTRNGTQLONFLDRSSVLDVGYSPNRRNEPLTGN | 710 |
| Db | 6900 | LRMTRNGTQLONFLDRSSVLDVGYSPNRRNEPLTGN | 6936 |

| | |
|----------|------------------------------------|
| RESULT 4 | |
| AAB12554 | |
| ID | AAB12554 standard; protein; 833 AA |

AC AAB12554

DT 07-NOV-2000 (first entry)

DE Human ovarian carcinoma antigen O772P protein SEQ ID NO:389.

KW Human; ovarian carcinoma; therapy; diagnosis;
KW tumour antigen; identification; cytostatic; gene therapy; vaccine

OS Homo sapiens.

PN WO2000036107-A2

PD 22-JUN-2000

PF 17-DEC-1999; 99WO-US030270.

| | | |
|-------------|--------------|----------------|
| PR | 17-DEC-1998; | 98US-00215681. |
| 17 DEC 1998 | | 98US-00215603 |

PR 23-JUN-1999; 99US-00338933;
24-SEP-1999; 99IS-00404879

XX
PA
(COBT-) COBTYA CORP -

XX
PT Mircham JL. King GE. Algate

WPI: 2000-431589/37

DN N-FOOD, 444/00/0:
XX

PT encoding it, useful

XX

XX

CC Immunogenic portion of an ovarian carcinoma

cytostatic activity and can be used in gene therapy and vaccines. Ova-

CC For the prevention, diagnosis and treatment of cancer, preferably ovarian

ovarian carcinoma polynucleotides and p

Sequence 833 AA;

| | | | | |
|---------------------------|-------|--------------------|----------|------------|
| Query Match | 92.9% | Score 3663 | DB 3 | Length 833 |
| Best Local Similarity | 99.9% | Pred. No. 2,3e-315 | | |
| Matches 696; Conservative | 0 | Mismatches 1 | Indels 0 | Gaps 0 |

QY 14 FTHRSSVSTSTSGTGPVVLGASKTPASIFGSSAASHLLILEFTLNPFITTLNLYEENMMPG 73

Db 78 FTHRSSVSTSTSGTGPVVLGASKTPASIFGSSAASHLLILEFTLNPFITTLNLYEENMMPG 137


```
QY 74 SRKFNTERVLQGLRLPLFKNTSVGPLYSGCRLLTLRPEKDGATGDAICTHRPDPTGP 133
DB 138 SRKFNTERVLQGLRLPLFKNTSVGPLYSGCRLLTLRPEKDGATGDAICTHRPDPTGP 197
QY 134 GLDREQLYELSQLTHSITELGPTTLDRDSL YVNGFTHRSSVPTTSGVSEEPFLNFT 193
DB 198 GLDREQLYELSQLTHSITELGPTTLDRDSL YVNGFTHRSSVPTTSGVSEEPFLNFT 257
QY 194 INNLRYADMGGQPSLKFENITDVMKHLSPLEFORSSLGARYGCRVIALRSYKNGAETR 253
DB 258 INNLRYADMGGQPSLKFENITDVMKHLSPLEFORSSLGARYGCRVIALRSYKNGAETR 317
QY 254 VDLCTYLQPLSGPGLPIKQVPHLSQOHTGIRLGYSDYDKOSLYLVNGYNEGPDPPT 313
DB 318 VDLCTYLQPLSGPGLPIKQVPHLSQOHTGIRLGYSDYDKOSLYLVNGYNEGPDPPT 377
QY 314 TPXPATTFPLPSEATTAMGYHLKTLTLNFTISNLQYSPDMGKSATFNSTEGVLOHLR 373
DB 378 TPXPATTFPLPSEATTAMGYHLKTLTLNFTISNLQYSPDMGKSATFNSTEGVLOHLR 437
QY 374 PLFOKSMGPFYLGCOLISLRPEKDGATGVDTTCTYHPDPVPGGLDIQOLYWELSQLT 433
DB 438 PLFOKSMGPFYLGCOLISLRPEKDGATGVDTTCTYHPDPVPGGLDIQOLYWELSQLT 497
QY 434 GVTOLGFYVLDRLSLFINGYAPONLSIRGEYOINFHIVMNLSPNPTSEYITLLRDIQ 493
DB 498 GVTOLGFYVLDRLSLFINGYAPONLSIRGEYOINFHIVMNLSPNPTSEYITLLRDIQ 557
QY 494 DKVTLTYKGSQLDHTRFCLVTLNLTMDSVLVTKALFSSNLDPSLVEQVFLDKTLNASH 553
DB 558 DKVTLTYKGSQLDHTRFCLVTLNLTMDSVLVTKALFSSNLDPSLVEQVFLDKTLNASH 617
QY 554 WLGSTYQLVDIHVTEMESVYOPTSSSTQHFYLNFTTNLPYSQDKAOPGTNNYQNRK 613
DB 618 WLGSTYQLVDIHVTEMESVYOPTSSSTQHFYLNFTTNLPYSQDKAOPGTNNYQNRK 677
QY 614 NIEDALNQLFRNSSIKSYFSDCOVSTFRSVPNRHHTGVDSL CNFSLARRVDVAIYEER 673
DB 678 NIEDALNQLFRNSSIKSYFSDCOVSTFRSVPNRHHTGVDSL CNFSLARRVDVAIYEER 737
QY 674 LRMTNNGTOLQNFITLDRSSVLVDGYSPPNRNEPLTGN 710
DB 738 LRMTNNGTOLQNFITLDRSSVLVDGYSPPNRNEPLTGN 774

RESULT 5
ABP30898
ID ABP30898 standard; protein, 833 AA.
XX
AC ABP30898;
XX
DB 02-JUL-2002 (first entry)
XX
DE 0772P clone 21003.
XX
KM Human; immunostimulant; cytostatic; cancer; ovarian carcinoma.
XX
OS Homo sapiens.
XX
PN WO200206317-A2.
XX
PD 24-JUN-2002.
XX
PF 17-JUL-2001; 2001WO-US022635.
XX
PR 17-JUL-2000; 2000US-00617747.
PR 10-SEP-2000; 2000US-00636801.
PR 20-SEP-2000; 2000US-0067857.
PR 04-APR-2001; 2001US-00827271.
PR 18-JUN-2001; 2001US-00884441.
XX
XX (CORI-) CORIXA CORP.
XX
```

```
PI Mitcham JL, King GE, Algate PA, Fling SP, Retter MW, Fanger GR;
PI Reed SG, Vedvick TS, Carter D, Hill P, Albone E;
XX WPI; 2002-164781/21.
XX DR N-PSDB; ABN72969.
XX
PT Polypeptides comprising an immunogenic portion of an ovarian carcinoma
PT patient or its variants, useful for stimulating an immune response in a
PT patient and treating ovarian cancer.
XX
XX Example 2; Page 316-318; 408pp; English.
XX
CC This invention relates to polypeptides comprising an immunogenic portion
CC of an ovarian carcinoma protein which access an immunostimulant and is
CC cytostatic. The polypeptides, polynucleotides, antibodies, fusion
CC proteins, T cell populations and antineoplastic cells that express
CC the polypeptides are useful for stimulating an immune response in a
CC patient and treating ovarian cancer. This sequence represents protein
CC related to the invention
XX
SQ Sequence 833 AA;
XX
Query Match 92.9%; Score 3663; DB 5; Length 833;
Best local Similarity 99.9%; Pred. No. 2,3e-315;
Matches 696; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 14 PTHRSSVSTTSPGTPYVLGASKTPASIFGSAASHLLILFTLNTTNLRYYENMWG 73
DB 78 PTHRSSVSTTSPGTPYVLGASKTPASIFGSAASHLLILFTLNTTNLRYYENMWG 137
QY 74 SRKFNTERVLQGLRLPLFKNTSVGPLYSGCRLLTLRPEKDGATGDAICTHRPDPTGP 133
DB 138 SRKFNTERVLQGLRLPLFKNTSVGPLYSGCRLLTLRPEKDGATGDAICTHRPDPTGP 197
QY 134 GLDREQLYELSQLTHSITELGPTTLDRDSL YVNGFTHRSSVPTTSGVSEEPFLNFT 193
DB 198 GLDREQLYELSQLTHSITELGPTTLDRDSL YVNGFTHRSSVPTTSGVSEEPFLNFT 257
QY 194 INNLRYADMGGQPSLKFENITDVMKHLSPLEFORSSLGARYGCRVIALRSYKNGAETR 253
DB 258 INNLRYADMGGQPSLKFENITDVMKHLSPLEFORSSLGARYGCRVIALRSYKNGAETR 317
QY 254 WLGSTYQLVDIHVTEMESVYOPTSSSTQHFYLNFTTNLPYSQDKAOPGTNNYQNRK 613
DB 618 WLGSTYQLVDIHVTEMESVYOPTSSSTQHFYLNFTTNLPYSQDKAOPGTNNYQNRK 677
QY 614 NIEDALNQLFRNSSIKSYFSDCOVSTFRSVPNRHHTGVDSL CNFSLARRVDVAIYEER 673
DB 678 NIEDALNQLFRNSSIKSYFSDCOVSTFRSVPNRHHTGVDSL CNFSLARRVDVAIYEER 737
QY 674 LRMTNNGTOLQNFITLDRSSVLVDGYSPPNRNEPLTGN 710
DB 738 LRMTNNGTOLQNFITLDRSSVLVDGYSPPNRNEPLTGN 774
```

```

RESULT 6
ADA08542
ID ADA08542 standard; protein; 833 AA.
XX
XX ADA08542;
XX
XX 06-NOV-2003 (first entry)
XX
XX Human ovarian carcinoma antigen O772P #2.
XX
XX human; gene therapy; ovarian cancer; cancer.
XX
XX Homo sapiens.
XX
XX US2003091580-A1.
XX
XX 15-MAY-2003.
XX
XX 17-JUL-2001; 2001US-00907969.
XX
XX 18-JUN-2001; 2001US-00884441.
XX
XX (MITC/) MITCHAM J L.
XX
XX (KING/) KING G E.
XX
XX (ALGA/) ALGATE P A.
XX
XX (FLIN/) FLING S P.
XX
XX (RETT/) RETTER M W.
XX
XX (FANG/) FANGER G R.
XX
XX (REED/) REED S G.
XX
XX (VEDV/) VEDVICK T S.
XX
XX (CART/) CARTER D.
XX
XX (HILL/) HILL P.
XX
XX (ALBO/) ALBONE E.
XX
XX Mitcham JL, King GE, Algate PA, Fling SP, Retter MW, Fanger GR;
XX
XX Reed SG, Vedvick TS, Carter D, Hill P, Albone E;
XX
XX WPI; 2003-532352/50.
XX
XX New isolated O772P polypeptides and polynucleotides, useful in gene
XX
XX therapy, particularly for treating or diagnosing cancer, e.g. ovarian
XX
XX cancer.
XX
XX Example 2; SEQ ID NO 389; 371bp; English.
XX
XX The invention relates to an isolated O772P polypeptide, which has the
XX
XX structure fully defined in the specification. The composition containing
XX
XX the polypeptides, polynucleotides, fusion proteins, antibodies, T-cells
XX
XX or antigen presenting cells are useful for stimulating an immune response
XX
XX and treating ovarian cancer. Detecting the presence of the
XX
XX polynucleotides and polypeptides is useful for diagnosing cancer. Ovarian
XX
XX carcinoma cDNAs and protein cDNAs were identified using microarray
XX
XX technology. The present sequence represents a human ovarian carcinoma
XX
XX antigen.
XX
XX Sequence 833 AA;
SQ
Query Match 92.9%; Score 3663; DB 7; Length 833;
Best Local Similarity 99.9%; Pred. No. 2,3e-315;
Matches 696; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 14 FTHRSSYSTSTPTPTVYLGASTKTPASISGSAASHLLIFLTNFTITNLRYENMMPG 73
DB 78 FTHRSSVSTSTPTPTVYLGASTKTPASISGSAASHLLIFLTNFTITNLRYENMMPG 137
OY 74 SRKNTTERVYQGLRPLFKNTSVGPLYSGRLTLRLPEKDGKATGVDAICTHRPDPGP 133
DB 138 SRKNTTERVYQGLRPLFKNTSVGPLYSGRLTLRLPEKDGKATGVDAICTHRPDPGP 197
OY 134 GLDREQLYLELSQLTHTSITELGPTTLDRDSLRYNGFTHRSSVPTTSGVSEEPFTLNFT 193

```

```

DB 198 GLDREQLYLELSQLTHTSITELGPTTLDRDSLRYNGFTHRSSVPTTSGVSEEPFTLNFT 257
OY 194 INNLRVADMGQSGSLKFNITTDNMKHLSPLEFORSISGARYTCRYALRSVNGAETR 253
DB 258 INNLRVADMGQSGSLKFNITTDNMKHLSPLEFORSISGARYTCRYALRSVNGAETR 317
OY 254 VDLICTYLOPLSGPGLPIKQVFNHLSQOTHTGITELGPISLDKSLYUNGNEPQDPDEPT 313
DB 318 VDLICTYLOPLSGPGLPIKQVFNHLSQOTHTGITELGPISLDKSLYUNGNEPQDPDEPT 377
OY 314 TPKPATTFELPPLSATTAMGYHLKTLNFTISNLQVSPDMKGSATFNTSTGVLQHLR 373
DB 378 TPKPATTFELPPLSATTAMGYHLKTLNFTISNLQVSPDMKGSATFNTSTGVLQHLR 437
OY 374 PLFQKSSMGPPYLCCQLISLRPEKDGATGVDTCTHPPDPVPGGLDIQQLYMLSQLTH 433
DB 438 PLFQKSSMGPPYLCCQLISLRPEKDGATGVDTCTHPPDPVPGGLDIQQLYMLSQLTH 497
OY 434 GVTOLGFYVLDRLSLFINGVAPONLSIRGEYQINPHIVNMNLSNPDPTSSSEYITLLRDIQ 493
DB 498 GVTOLGFYVLDRLSLFINGVAPONLSIRGEYQINPHIVNMNLSNPDPTSSSEYITLLRDIQ 557
OY 494 DKVYTLVYKSQLHDTFRFCVLTNLTMDSVLYTVKALFSSNLDPSLYEGVFLDKTLNLSFH 553
DB 558 DKVYTLVYKSQLHDTFRFCVLTNLTMDSVLYTVKALFSSNLDPSLYEGVFLDKTLNLSFH 617
OY 554 WLGSTYQVLDIHVTEMESVYOPTSSSTQHFYLNFTTNLPYSODKAOPTTYNQNRK 613
DB 618 WLGSTYQVLDIHVTEMESVYOPTSSSTQHFYLNFTTNLPYSODKAOPTTYNQNRK 677
OY 614 NIEDALNQLFRNSIKSYFSDCOVSTFRSVPRRHHTGVDSLGNFGLARRVDRAVAYEEF 673
DB 678 NIEDALNQLFRNSIKSYFSDCOVSTFRSVPRRHHTGVDSLGNFGLARRVDRAVAYEEF 737
OY 674 LEMTRNGTOLQNFITLDRSSVLDVGYSPNRNEPTLNS 710
DB 738 LEMTRNGTOLQNFITLDRSSVLDVGYSPNRNEPTLNS 774
RESULT 7
ADF08885
ID ADF08885 standard; protein; 833 AA.
XX
XX ADF08885;
XX
XX 12-FEB-2004 (first entry)
XX
XX Secreted ovarian carcinoma antigen seq'd 389.
XX
XX gene therapy; protein therapy; vaccine; antibody inhibition;
XX
XX breast cancer; restorative therapy; diagnostic agent; immunoassay;
XX
XX secreted ovarian carcinoma antigen.
XX
XX Homo sapiens.
XX
XX US2003124140-A1.
XX
XX 03-JUL-2003.
XX
XX 17-JUL-2002; 2002US-00198053.
XX
XX 17-DEC-1998; 98US-00215681.
XX
XX 17-DEC-1998; 98US-00216003.
XX
XX 23-JUN-1999; 99US-00338933.
XX
XX 24-SEP-1999; 99US-00404879.
XX
XX 17-JUL-2000; 2000US-00617747.
XX
XX 10-AUG-2000; 2000US-00636801.
XX
XX 20-SEP-2000; 2000US-00667857.
XX
XX 04-APR-2001; 2001US-00827271.
XX
XX 18-JUN-2001; 2001US-00884441.
XX
XX 17-JUL-2001; 2001US-00907969.
XX
XX (CORI-) CORIXA CORP.

```

XX
PI Bangur GS, Retter MM, Fanger GR, Hill P;
XX
DR WPI; 2003-897152/82.
XX
DR N-PSDB; ADF08882.
XX
PT Oncogenic nucleic acids useful for the prevention, diagnosis and
XX treatment of breast cancer.
XX
PS Example 2; SEQ ID NO 389; 399pp; English.
XX
CC The invention describes nucleic acids (I) and the polypeptides (II) they
CC encode. The nucleic acids (I) may be used for preventing, diagnosing and
CC treating diseases related to their aberrant expression i.e. breast
CC cancers. For example, (I) and (II) may be used to treat disorders
CC associated with decreased expression by rectifying mutations or deletions
CC in a patient's genome that affect the activity of (II) by expressing
CC inactive proteins or to supplement the patients own production of (II).
CC Additionally, (I) may be used to produce (II), by inserting (I) into a
CC host cell and culturing the cell to express the protein (II). (I) And its
CC complementary sequences may also be used as DNA probes in diagnostic
CC assays to detect and quantitate the presence of similar nucleic acids in
CC samples, and therefore which patients may be in need of restorative
CC therapy. The host cell may also be used as antigens in the production of
CC antibodies against (II) and in assays to identify modulators of (II)'s
CC expression and activity. The anti-(II) antibodies, agonists and
CC antagonists may be used to regulate expression and activity and as
CC diagnostic agents for detecting the presence of (II) in samples (e.g. by
CC immunoassay). This sequence represents a secreted ovarian carcinoma
CC antigen.
XX
XX
SQ Sequence 833 AA;

Query Match 92.9%; Score 3663; DB 7; Length 833;
Best Local Similarity 99.9%; Pred. No. 2.3e-315;
Matches 696; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 14 FTHRSSVSTSTPTGTPVYLGASTPASTFGPSAASHLLILFTLNTTINLRYEENMWP 73
DB 78 FTHRSSVSTSTPTGTPVYLGASTPASTFGPSAASHLLILFTLNTTINLRYEENMWP 137
QY 74 SRKFNTERVLOGLRPLFKNTSVGPLYSGCRLTLRPEKGEATGVDAICTHRPDPTGP 133
DB 138 SRKFNTERVLOGLRPLFKNTSVGPLYSGCRLTLRPEKGEATGVDAICTHRPDPTGP 197
QY 134 GLDREQLYLELSQTLHSITELGPTLDRDSL YVNGFTHRSSVPTSTGVSEEPFTLNT 193
DB 198 GLDREQLYLELSQTLHSITELGPTLDRDSL YVNGFTHRSSVPTSTGVSEEPFTLNT 257
QY 194 INNLRVADMGQPGSLKFNITDVMKILSPFORSSLGARYTCRVIALRSVNGAETR 253
DB 258 INNLRVADMGQPGSLKFNITDVMKILSPFORSSLGARYTCRVIALRSVNGAETR 317
QY 254 VDLICTYLOPLSGRPIKQVFEHLISQOTGIRLGYSLDKOSLYVNGNEGPDEPPT 313
DB 318 VDLICTYLOPLSGRPIKQVFEHLISQOTGIRLGYSLDKOSLYVNGNEGPDEPPT 377
QY 314 TPKPATFTLPLSEATTAMGVHLKTLNFTTISNLQYSPMGKGSATFNSTEGVLQHLR 373
DB 378 TPKPATFTLPLSEATTAMGVHLKTLNFTTISNLQYSPMGKGSATFNSTEGVLQHLR 437
QY 374 PLFOKSMGPFYLGCOQLISRPKDGATGVDITCTYHPDPVGGDLIDQLYWEISQTLH 433
DB 438 PLFOKSMGPFYLGCOQLISRPKDGATGVDITCTYHPDPVGGDLIDQLYWEISQTLH 497
QY 434 GVTOLGPFYVLDRLSTFNNGVAPONLSTRGEVQINFNHVMNNLNPDPSTSEYITLARDIQ 493
DB 498 GVTOLGPFYVLDRLSTFNNGVAPONLSTRGEVQINFNHVMNNLNPDPSTSEYITLARDIQ 557
QY 494 DKYTTLYKGSQLDHDFRCLVTNLTMDSVLTVKALFPSSNLDPSLVEQVLDLTNLASFH 553
DB 558 DKYTTLYKGSQLDHDFRCLVTNLTMDSVLTVKALFPSSNLDPSLVEQVLDLTNLASFH 617

QY 554 WLGSTYOLVDIHTEMESSYOPTSSSSSTOHFLNTFTINLPSYOPKAPGPTTNGRNR 613
DB 618 WLGSTYOLVDIHTEMESSYOPTSSSSSTOHFLNTFTINLPSYOPKAPGPTTNGRNR 677
QY 614 NIEDALNOLFNRNSISYFSDCCVSTFRSVPNRHGTGVDSLCNFSPLARRDVAIYE 673
DB 678 NIEDALNOLFNRNSISYFSDCCVSTFRSVPNRHGTGVDSLCNFSPLARRDVAIYE 737
QY 674 LRMTNGTOLQNTLDRSSVLDVGYSPNRPETLNGS 710
DB 738 LRMTNGTOLQNTLDRSSVLDVGYSPNRPETLNGS 774

RESULT 8
ID ADG46172 standard; protein; 833 AA.
AC ADG46172;
DT 26-FEB-2004 (first entry)
XX
XX
DE Human ovarian carcinoma polypeptide #3.
XX
XX
KM Human; ovarian carcinoma; OBE; ovarian cancer; secreted tumour antigen;
KM cytostatic; O772P.
OS Homo sapiens.
OS
PN US2003165504-A1.
XX
XX
PD 04-SEP-2003.
XX
XX
PF 04-APR-2001; 2001US-00827271.
XX
XX
PR 17-DEC-1998; 98US-00215681.
PR 17-DEC-1998; 98US-00216003.
PR 23-JUN-1999; 99US-00338933.
PR 24-SEP-1999; 99US-00404879.
PR 17-JUL-2000; 2000US-00617747.
PR 10-AUG-2000; 2000US-00636801.
PR 20-SEP-2000; 2000US-00667857.
XX
XX
PA (RETT/) RETTER M W.
PA (FANG/) FANGER G R.
XX
XX
PI Retter MM, Fanger GR;
XX
XX
DR WPI; 2003-898035/82.
XX
XX
PT New isolated OBE or O772P polypeptides, useful for diagnosing,
PT preventing, treating and monitoring cancer, e.g. ovarian cancer,
PT stimulating the immune response in patient.
XX
XX
PS Claim 5; SEQ ID NO 389; 290pp; English.
XX
XX
CC The invention relates to human ovarian carcinoma polypeptides, designated
CC OBE or O772P, and the polynucleotides encoding them. The invention also
CC relates to methods for inhibiting the development of cancer, e.g. ovarian
CC cancer in a patient, methods for stimulating and/or expanding T cells and
CC methods for identifying secreted tumour antigens. The polypeptides,
CC compositions, antibodies to the polypeptides and methods are useful for
CC diagnosing, preventing, treating and monitoring cancer, e.g. ovarian
CC cancer. The composition is particularly useful for stimulating an immune
CC response in patient. This sequence represents a human ovarian carcinoma
CC polypeptide of the invention.
XX
XX
SQ Sequence 833 AA;

Query Match 92.9%; Score 3663; DB 7; Length 833;
Best Local Similarity 99.9%; Pred. No. 2.3e-315;
Matches 696; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 14 FTHRSSVSTSTPTGTPVYLGASTPASTFGPSAASHLLILFTLNTTINLRYEENMWP 73

```

Db      78 FTHRSSVSTSTPGPTVYLGAASKTPASIFGSAASHLLILFTLNFTINLRYEENMPG 137
Qy      74 SRKFNTTTERVLOGLRLPLFKNTSVGPLYSGCRLLTLRPRKDEAGTVDAILCHRPDPPTG 133
Db      138 SRKFNTTTERVLOGLRLPLFKNTSVGPLYSGCRLLTLRPRKDEAGTVDAILCHRPDPPTG 197
Qy      134 GUDREOLYLELSQTLHSITELGPTLDRDSLVLVNGGTHHSSVPTTSTGVVSEEPFLNFT 193
Db      198 GUDREOLYLELSQTLHSITELGPTLDRDSLVLVNGGTHHSSVPTTSTGVVSEEPFLNFT 257
Qy      194 INNLRYMADMGQGSILKFNITDVMGHLSPLEFORSISLARVTGCRVIALRSVKNGAETR 253
Db      258 INNLRYMADMGQGSILKFNITDVMGHLSPLEFORSISLARVTGCRVIALRSVKNGAETR 317
Qy      254 VDLCTYLOPLSGPGRLPIKQVFHELSQOTHGITRLGRLPSLDKDSLVLNGYNEGPDEPPT 313
Db      318 VDLCTYLOPLSGPGRLPIKQVFHELSQOTHGITRLGRLPSLDKDSLVLNGYNEGPDEPPT 377
Qy      314 TRKPATTFPLPSEATTAMGYHLKTLTNFTISNLQYSPDMGKSATFNSSTEGVLOHLRL 373
Db      378 TRKPATTFPLPSEATTAMGYHLKTLTNFTISNLQYSPDMGKSATFNSSTEGVLOHLRL 437
Qy      374 PLFOKSSMGPFYLGCOLISLRPKDGAATGVDTCYHDPVPGRLDIQOLYWELSQTLH 433
Db      438 PLFOKSSMGPFYLGCOLISLRPKDGAATGVDTCYHDPVPGRLDIQOLYWELSQTLH 497
Qy      434 GTVQLGFYVLDRLDSLFLNGYAPONLSIRGEYQINFHIVMNLNPNPTSEYITLLRDIO 493
Db      498 GTVQLGFYVLDRLDSLFLNGYAPONLSIRGEYQINFHIVMNLNPNPTSEYITLLRDIO 557
Qy      494 DKVTTLTKSGQLHDTFRFLCVTNLTMDSVLVYTKALFSSSLDLSLYEQVFLDITLNASFH 553
Db      558 DKVTTLTKSGQLHDTFRFLCVTNLTMDSVLVYTKALFSSSLDLSLYEQVFLDITLNASFH 617
Qy      554 WLGSTYQLVLDIHTEMESSYVQPTSSSTQHFYLNFTITNLPSODKAOPGTITNYQRNKR 613
Db      618 WLGSTYQLVLDIHTEMESSYVQPTSSSTQHFYLNFTITNLPSODKAOPGTITNYQRNKR 677
Qy      614 NIEDALNQLFRNSSIKSYFSDCCVSTFRSVVNRHHTGVDSLGNFSPLARVRDVAIYEER 673
Db      678 NIEDALNQLFRNSSIKSYFSDCCVSTFRSVVNRHHTGVDSLGNFSPLARVRDVAIYEER 737
Qy      674 LRMTNRGTQLONTFLDRSSVLVDGYSPNNEPITLGN 710
Db      738 LRMTNRGTQLONTFLDRSSVLVDGYSPNNEPITLGN 774

```

```

PR      10-AUG-2000; 2000US-00636801.
PR      20-SEP-2000; 2000US-00667857.
PR      04-APR-2001; 2001US-00827271.
PR      18-JUN-2001; 2001US-00884441.
PR      17-JUL-2001; 2001US-00907969.
PR      17-JUL-2002; 2002US-00198053.
PA      (CORI-) CORIXA CORP.
XX      PI      Bangur CS, Retter MW, Fanger GR, Hill P;
XX      WI1; 2005-151645/16.
XX      DR      WPI; 2005-151645/16.
XX      PT      New ovarian carcinoma polynucleotides, preferably cDNAs, useful for
XX      PT      diagnosing, preventing and treating diseases, such as ovarian cancer, and
XX      PS      for eliciting humoral and/or cellular immune response.
XX      Example 2; SEQ ID NO 389; 398bp; English.
XX      CC      The invention relates to an isolated polynucleotide comprising a sequence
XX      CC      of, a sequence hybridizing under highly stringent conditions to, or
XX      CC      having at least 75, or 90% identity to a 849 or 1399 bp, fully defined in
XX      CC      the specification, its complement or degenerate variants, or a sequence
XX      CC      of at least 20 contiguous residues of the 849 or 1399 bp sequence. The
XX      CC      polynucleotides and polypeptides are useful for diagnosing, preventing
XX      CC      and/or treating diseases, such as ovarian cancer, and for eliciting humoral
XX      CC      carcinoma antigen immune response. This sequence corresponds to an ovarian
XX      CC      carcinoma antigen protein of the invention.
XX      SQ      Sequence 833 AA;
Qy      Query Match      92.9%; Score 3663; DB 9; Length 833;
Qy      Best Local Similarity 99.9%; Pred. No. 2,3e-315;
Qy      Matches 696; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Db      14 FTHRSSVSTSTPGPTVYLGAASKTPASIFGSAASHLLILFTLNFTINLRYEENMPG 73
Db      78 FTHRSSVSTSTPGPTVYLGAASKTPASIFGSAASHLLILFTLNFTINLRYEENMPG 137
Qy      74 SRKFNTTTERVLOGLRLPLFKNTSVGPLYSGCRLLTLRPRKDEAGTVDAILCHRPDPPTG 133
Db      138 SRKFNTTTERVLOGLRLPLFKNTSVGPLYSGCRLLTLRPRKDEAGTVDAILCHRPDPPTG 197
Qy      134 GUDREOLYLELSQTLHSITELGPTLDRDSLVLVNGGTHHSSVPTTSTGVVSEEPFLNFT 193
Db      198 GUDREOLYLELSQTLHSITELGPTLDRDSLVLVNGGTHHSSVPTTSTGVVSEEPFLNFT 257
Qy      194 INNLRYMADMGQGSILKFNITDVMGHLSPLEFORSISLARVTGCRVIALRSVKNGAETR 253
Db      258 INNLRYMADMGQGSILKFNITDVMGHLSPLEFORSISLARVTGCRVIALRSVKNGAETR 317
Qy      254 VDLCTYLOPLSGPGRLPIKQVFHELSQOTHGITRLGRLPSLDKDSLVLNGYNEGPDEPPT 313
Db      318 VDLCTYLOPLSGPGRLPIKQVFHELSQOTHGITRLGRLPSLDKDSLVLNGYNEGPDEPPT 377
Qy      314 TRKPATTFPLPSEATTAMGYHLKTLTNFTISNLQYSPDMGKSATFNSSTEGVLOHLRL 373
Db      378 TRKPATTFPLPSEATTAMGYHLKTLTNFTISNLQYSPDMGKSATFNSSTEGVLOHLRL 437
Qy      374 PLFOKSSMGPFYLGCOLISLRPKDGAATGVDTCYHDPVPGRLDIQOLYWELSQTLH 433
Db      438 PLFOKSSMGPFYLGCOLISLRPKDGAATGVDTCYHDPVPGRLDIQOLYWELSQTLH 497
Qy      434 GTVQLGFYVLDRLDSLFLNGYAPONLSIRGEYQINFHIVMNLNPNPTSEYITLLRDIO 493
Db      498 GTVQLGFYVLDRLDSLFLNGYAPONLSIRGEYQINFHIVMNLNPNPTSEYITLLRDIO 557
Qy      494 DKVTTLTKSGQLHDTFRFLCVTNLTMDSVLVYTKALFSSSLDLSLYEQVFLDITLNASFH 553
Db      558 DKVTTLTKSGQLHDTFRFLCVTNLTMDSVLVYTKALFSSSLDLSLYEQVFLDITLNASFH 617
Qy      554 WLGSTYQLVLDIHTEMESSYVQPTSSSTQHFYLNFTITNLPSODKAOPGTITNYQRNKR 613

```

Db 618 WLGSYQLVDIHVTEMESSVYOPTSSSTQHFLNFTITNLPSQDKAOPGTTNYQRNK 677

QY 614 NIEDALNQLFRNSSIKSYFSDCOVSTFRSVPNRHHTGVDSLGNFSPILARRVDAIYEER 673

Db 678 NIEDALNQLFRNSSIKSYFSDCOVSTFRSVPNRHHTGVDSLGNFSPILARRVDAIYEER 737

QY 674 LRMTRNGTOLQNFLLDRSSVLYDGYSPNNEPLTGN 710

Db 738 LRMTRNGTOLQNFLLDRSSVLYDGYSPNNEPLTGN 774

RESULT 10

AA12552

ID AA12552 standard; protein; 914 AA.

AC AA12552;

XX 07-NOV-2000 (first entry)

XX 07-NOV-2000 (first entry)

DE Human ovarian carcinoma antigen O772P protein SEQ ID NO:312.

XX

KW Human; ovarian carcinoma; ovarian cancer; therapy; diagnosis;

KM tumour antigen; identification; cytostatic; gene therapy; vaccine.

XX

OS Homo sapiens.

XX

PN MO200036107-A2.

XX

PD 22-JUN-2000.

XX

PF 17-DEC-1999; 99MO-US030270.

XX

PR 17-DEC-1998; 98US-00215681.

PR 17-DEC-1998; 98US-00216003.

PR 23-JUN-1999; 99US-00338933.

PR 24-SEP-1999; 99US-00404879.

XX

PA (CORI-) CORIXA CORP.

PI Mitcham JL, King GE, Aigate PA, Frudakis TN;

XX

DR WPI; 2000-431589/37.

DR N-PSDB; AAAT0001.

XX

PT Immunogenic portion of an ovarian carcinoma protein and the nucleic acid

PT encoding it, useful for the diagnosis, prevention and treatment of

PT cancer, preferably ovarian cancer.

XX

PS Example 2; Page 172-174; 299pp; English.

XX

CC The present invention describes an isolated polypeptide comprising an

CC immunogenic portion of an ovarian carcinoma protein (or its variants).

CC Ovarian carcinoma proteins, and polynucleotides encoding them, have

CC cytostatic activity and can be used in gene therapy and vaccines. Ovarian

CC carcinoma polypeptides, nucleic acids, antibodies and vaccines are useful

CC for the prevention, diagnosis and treatment of cancer, preferably ovarian

CC cancer. AA469691 to AA470077 and AA12552 to AA12557 represent human

CC ovarian carcinoma polynucleotides and proteins used in the

CC exemplification of the present invention

XX

SO Sequence 914 AA;

Query Match 92.9%; Score 3663; DB 3; Length 914;

Best Local Similarity 99.9%; Pred. No. 2,7e-315;

Matches 696; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 14 FTHRSSVSTSTGTGTIVYIGAKTPASTIGPSAASHLLFLFTNLFTITNLRYEENMWG 73

Db 159 FTHRSSVSTSTGTGTIVYIGAKTPASTIGPSAASHLLFLFTNLFTITNLRYEENMWG 218

QY 74 SRKFTTERVLOGILRLPLFKNTSVGPLYSGCRLTLRLPEKOGATGVDAICTHRPDPGP 133

Db 219 SRKFTTERVLOGILRLPLFKNTSVGPLYSGCRLTLRLPEKOGATGVDAICTHRPDPGP 278

QY 134 GIDREOLYLIELSOLTHSITELEGPYTLDRDSL.YVNGPTHRSSVPTTSAGVSEEPPTLNFT 193

Db 279 GIDREOLYLIELSOLTHSITELEGPYTLDRDSL.YVNGPTHRSSVPTTSAGVSEEPPTLNFT 338

QY 194 INNLRYMADMGPGSLKFNITDNVMGHLSPLEFORSSLGARYTGCRVIALRSYKNGAETR 253

Db 339 INNLRYMADMGPGSLKFNITDNVMGHLSPLEFORSSLGARYTGCRVIALRSYKNGAETR 398

QY 254 VDLCTYLOPLSGRGLPIKQVTHELSQOHTGITRLGPIYLDKDSL.YLNGNEGPDPPT 313

Db 399 VDLCTYLOPLSGRGLPIKQVTHELSQOHTGITRLGPIYLDKDSL.YLNGNEGPDPPT 458

QY 314 TPKPATFPLPLSEATAMGYHLKTLTLNFTISNLQYSPDMGKSAFENSTEGVLOHLR 373

Db 459 TPKPATFPLPLSEATAMGYHLKTLTLNFTISNLQYSPDMGKSAFENSTEGVLOHLR 518

QY 374 PLFOKSSMGPFYLGCOLISLRPEKGAATGVDTTCYHPDPVGPGLDIQOLYWELSQLT 433

Db 519 PLFOKSSMGPFYLGCOLISLRPEKGAATGVDTTCYHPDPVGPGLDIQOLYWELSQLT 578

QY 434 GVTOLGFVYLDRLDSEFINGYAPQNLISGEYOINPHIYVNMNLSNPDPTSSBYITLLRDIO 493

Db 579 GVTOLGFVYLDRLDSEFINGYAPQNLISGEYOINPHIYVNMNLSNPDPTSSBYITLLRDIO 638

QY 494 DKVTTLTKGSQLHDFRFLVNLNLTMDSVLYTVKALFSSNDPLSLVEQVFLDKTLNASH 553

Db 639 DKVTTLTKGSQLHDFRFLVNLNLTMDSVLYTVKALFSSNDPLSLVEQVFLDKTLNASH 698

QY 554 WLGSYQLVDIHVTEMESSVYOPTSSSTQHFLNFTITNLPSQDKAOPGTTNYQRNK 613

Db 699 WLGSYQLVDIHVTEMESSVYOPTSSSTQHFLNFTITNLPSQDKAOPGTTNYQRNK 758

QY 614 NIEDALNQLFRNSSIKSYFSDCOVSTFRSVPNRHHTGVDSLGNFSPILARRVDAIYEER 673

Db 759 NIEDALNQLFRNSSIKSYFSDCOVSTFRSVPNRHHTGVDSLGNFSPILARRVDAIYEER 818

QY 674 LRMTRNGTOLQNFLLDRSSVLYDGYSPNNEPLTGN 710

Db 819 LRMTRNGTOLQNFLLDRSSVLYDGYSPNNEPLTGN 855

RESULT 11

AA12552

ID AA12552 standard; protein; 914 AA.

AC AA12552;

XX 04-SEP-2001 (first entry)

XX

DE Human ovarian tumour-derived antigen O772P.

XX

KW Cytostatic; human; breast tumour protein; breast cancer; ovarian tumour;

KM antigen; O772P.

XX

OS Homo sapiens.

XX

PN MO200140269-A2.

XX

PD 07-JUN-2001.

XX

PF 29-NOV-2000; 2000MO-US032520.

XX

PR 30-NOV-1999; 99US-00451651.

PR 22-FEB-2000; 2000US-00510662.

PR 10-MAR-2000; 2000US-00523566.

PR 07-APR-2000; 2000US-00545066.

PR 15-MAY-2000; 2000US-00571025.

XX

PA (CORI-) CORIXA CORP.

PI Dillon DC, Day CH, Jiang Y, Houghton RL, Mitcham JL, Wang A;

DR WPI; 2001-356154/37.
DR N-PSDB; AAH55680.

XX Breast tumor polypeptides and the nucleic acids that encode them, useful
PT for the prevention, diagnosis and treatment of breast cancer.

XX Example 3; Page 187-188; 221pp; English.

XX The present invention relates to human breast tumour protein coding
CC sequences (see AAH55479-AAH55513, AAH55517-AAH55679 and AAH55682-
CC AAH55762). The breast tumour protein DNA sequences may be used in the
CC prevention, diagnosis and treatment of diseases associated with
CC inappropriate expression of the breast tumour protein e.g. breast cancer.
CC The present sequence is a human ovarian tumour-derived antigen, which was
CC used in an example from the present invention

XX Sequence 914 AA;

Query Match 92.9%; Score 3663; DB 4; Length 914;
Best Local Similarity 99.9%; Pred. No. 2,7e-315;
Matches 696; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 14 FTHRSSVSTSTTSGPTTVYLGASKTPASIFGSAASHLLILFTLNTTINLYEENMMRG 73
DB 159 FTHRSSVSTSTTSGPTTVYLGASKTPASIFGSAASHLLILFTLNTTINLYEENMMRG 218
QY 74 SRKFNTERVLOGLRLPLFKNTSVGPLYSGCRLLTLRPEKDGATGVDAICTHRPDPPTGP 133
DB 219 SRKFNTERVLOGLRLPLFKNTSVGPLYSGCRLLTLRPEKDGATGVDAICTHRPDPPTGP 278
QY 134 GIDREQLYLELSQTLHSITELGPLYTLDRDSLVLVNGFTHRSSVPTTSGVVSSEPTLNT 193
DB 279 GIDREQLYLELSQTLHSITELGPLYTLDRDSLVLVNGFTHRSSVPTTSGVVSSEPTLNT 338
QY 194 INNLRYMADMGOGSLKFNITDVMKHLSPLFORSISGARVYGCRIALRSYNGAETR 253
DB 339 INNLRYMADMGOGSLKFNITDVMKHLSPLFORSISGARVYGCRIALRSYNGAETR 398
QY 254 VDLCTYLOPLSGPGLPIKOVFHELSSQTHGIRLPGYSIDKSLVYNGNEBPDEPT 313
DB 399 VDLCTYLOPLSGPGLPIKOVFHELSSQTHGIRLPGYSIDKSLVYNGNEBPDEPT 458
QY 314 TPKPATTFLPLSEATTAMGYHLKTLTNTISNLQYSPMGKGSATFNSTEGVLOHLR 373
DB 459 TPKPATTFLPLSEATTAMGYHLKTLTNTISNLQYSPMGKGSATFNSTEGVLOHLR 518
QY 374 PLFOKSMGPFIYGCQOLISLRPEKDGATGVDTCTYHPDVGGLDIQOLYHELSQLTH 433
DB 519 PLFOKSMGPFIYGCQOLISLRPEKDGATGVDTCTYHPDVGGLDIQOLYHELSQLTH 578
QY 434 GVTOLGYYVLDRLSLFINGYAPQNLIRGEYQINFHIVNNLSNPDTSSSEYITLLRDQ 493
DB 579 GVTOLGYYVLDRLSLFINGYAPQNLIRGEYQINFHIVNNLSNPDTSSSEYITLLRDQ 638
QY 494 DKVTLLYKGSQLDHDTFFRCVLTNLTMDSVLYVYKALFSSNLDPSLVEQVFLDKTLNLSFH 553
DB 639 DKVTLLYKGSQLDHDTFFRCVLTNLTMDSVLYVYKALFSSNLDPSLVEQVFLDKTLNLSFH 698
QY 554 WLGSYLOLDIHTWEMESSYVQPTSSSTQHFIYNTITNLRYQODKAOQCTMYONRKR 613
DB 699 WLGSYLOLDIHTWEMESSYVQPTSSSTQHFIYNTITNLRYQODKAOQCTMYONRKR 758
QY 614 NIEBALNOLFNSSIKSYFSDCOVSTFRSVNRRHHTGVDSLGNFSLARVRAIYEEF 673
DB 759 NIEBALNOLFNSSIKSYFSDCOVSTFRSVNRRHHTGVDSLGNFSLARVRAIYEEF 818
QY 674 LRMTNQTOLQNTLDRSSVLDVGYSPNRNBLPTGNS 710
DB 819 LRMTNQTOLQNTLDRSSVLDVGYSPNRNBLPTGNS 855

```

RESULT 12
ABP30968

ID ABP30968 standard; protein; 914 AA.

XX AC ABP30968;

XX DT 02-JUN-2002 (first entry)

XX DE Hypothetical protein sequence for clone 0772P.

XX KM Human; immunostimulant; cytostatic; cancer; ovarian carcinoma.

XX OS Homo sapiens.

XX PN WO200206317-A2.

XX PD 24-JAN-2002.

XX PF 17-JUN-2001; 2001WO-US022635.

XX PR 17-JUN-2000; 2000US-00617747.

XX PR 10-AUG-2000; 2000US-00636801.

XX PR 20-SEP-2000; 2000US-00667857.

XX PR 04-APR-2001; 2001US-00827271.

XX PR 18-JUN-2001; 2001US-00884441.

XX PA (CORI-) CORIXA CORP.

XX PI Mitcham JU, King GE, Algate PA, Fling SP, Retter MW, Fanger GR,

XX PI Reed SG, Vedvick TS, Carter D, Hill F, Albone E;

XX DR WPI; 2002-164781/21.

XX DR N-PSDB; ABN72974.

XX PS Disclosure; Page 350-352; 408pp; English.

XX This invention relates to polypeptides comprising an immunogenic portion
CC of an ovarian carcinoma protein which acts as an immunostimulant and is
CC cytotoxic. The polypeptides, polynucleotides, antibodies, fusion
CC proteins, T cell populations and antigen presenting cells that express
CC the polypeptides are useful for stimulating an immune response in a
CC patient and treating ovarian cancer. This sequence represents protein
CC related to the invention

XX Sequence 914 AA;

Query Match 92.9%; Score 3663; DB 5; Length 914;
Best Local Similarity 99.9%; Pred. No. 2,7e-315;
Matches 696; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 14 FTHRSSVSTSTTSGPTTVYLGASKTPASIFGSAASHLLILFTLNTTINLYEENMMRG 73
DB 159 FTHRSSVSTSTTSGPTTVYLGASKTPASIFGSAASHLLILFTLNTTINLYEENMMRG 218
QY 74 SRKFNTERVLOGLRLPLFKNTSVGPLYSGCRLLTLRPEKDGATGVDAICTHRPDPPTGP 133
DB 219 SRKFNTERVLOGLRLPLFKNTSVGPLYSGCRLLTLRPEKDGATGVDAICTHRPDPPTGP 278
QY 134 GIDREQLYLELSQTLHSITELGPLYTLDRDSLVLVNGFTHRSSVPTTSGVVSSEPTLNT 193
DB 279 GIDREQLYLELSQTLHSITELGPLYTLDRDSLVLVNGFTHRSSVPTTSGVVSSEPTLNT 338
QY 194 INNLRYMADMGOGSLKFNITDVMKHLSPLFORSISGARVYGCRIALRSYNGAETR 253
DB 339 INNLRYMADMGOGSLKFNITDVMKHLSPLFORSISGARVYGCRIALRSYNGAETR 398
QY 254 VDLCTYLOPLSGPGLPIKOVFHELSSQTHGIRLPGYSIDKSLVYNGNEBPDEPT 313
DB 399 VDLCTYLOPLSGPGLPIKOVFHELSSQTHGIRLPGYSIDKSLVYNGNEBPDEPT 458
QY 314 TPKPATTFLPLSEATTAMGYHLKTLTNTISNLQYSPMGKGSATFNSTEGVLOHLR 373

```



```

Db      459 TPKATTFPLPSEATTAMGYHLKTLTLNFTISNLOYS PDMGKGSATFNSGEGVLQHLR 518
      374 PLFQKSMGPFYLGCOLISLRPEKDGATGVDITCTYHPDVGGLDIQQLYWELSQULTH 433
      519 PLFQKSMGPFYLGCOLISLRPEKDGATGVDITCTYHPDVGGLDIQQLYWELSQULTH 578
Qy      434 GVTQLGFFVLDROSLFINGYAPQNLSTIRGEYQINFNHVMNLSNPDPSTSSSEYITLLRDIQ 493
      579 GVTQLGFFVLDROSLFINGYAPQNLSTIRGEYQINFNHVMNLSNPDPSTSSSEYITLLRDIQ 638
Qy      494 DKVTLLYKGSQLDHTEFECVLTNLTMDSVLVTVYALFSSNLDPSLVEQVFLDKTLNASFH 553
      639 DKVTLLYKGSQLDHTEFECVLTNLTMDSVLVTVYALFSSNLDPSLVEQVFLDKTLNASFH 698
Db      554 WLGSSTYQLVLDIHVTEMESVYQPTSSSTQHFYINFTITNLPSQDKAQPCTTNYQGNKR 613
      699 WLGSSTYQLVLDIHVTEMESVYQPTSSSTQHFYINFTITNLPSQDKAQPCTTNYQGNKR 758
Qy      614 NIEDALNQLFRNSISIKYFSDCCVSTFRSVNRRHHTGVDSL CNFSPILARRVDAIYEEF 673
      759 NIEDALNQLFRNSISIKYFSDCCVSTFRSVNRRHHTGVDSL CNFSPILARRVDAIYEEF 818
Db      674 LRMTNQTQLQNFLLDRSSVLYDGYSPNREPLTGN 710
      819 LRMTNQTQLQNFLLDRSSVLYDGYSPNREPLTGN 855

```

RESULT 13

ABP30896 ID ABP30896 standard; protein; 914 AA.

AC ABP30896;

DT 02-JUL-2002 (first entry)

XX 0772P protein.

XX Human; immunostimulant; cytostatic; cancer; ovarian carcinoma.

OS Homo sapiens.

PN W0200206317-A2.

PD 24-JAN-2002.

PF 17-JUL-2001; 2001WO-US02635.

PR 17-JUL-2000; 2000US-00617747.

PR 10-AUG-2000; 2000US-00636801.

PR 20-SEP-2000; 2000US-00667857.

PR 04-APR-2001; 2001US-00827271.

PR 18-JUN-2001; 2001US-00884441.

XX (CORI-) CORIXA CORP.

PI Mitcham JL, King GE, Algate PA, Fling SP, Retter MW, Fanger GR;

PI Reed SG, Vedvick TS, Carter D, Hill P, Albone E;

XX WPI; 2002-164781/21.

XX N-PSDB; ABN72895.

XX Polypeptides comprising an immunogenic portion of an ovarian carcinoma

XX protein or its variants, useful for stimulating an immune response in a

XX patient and treating ovarian cancer.

XX Example 2; Page 291; 408bp; English.

```

CC patient and treating ovarian cancer. This sequence represents protein
CC related to the invention
XX      SQ      Sequence 914 AA;
      Query Match      92.9%; Score 3663; DB 5; Length 914;
      Best Local Similarity 99.9%; Pred. No. 2,7e-315;
      Matches 696; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy      14 FTHRSSVSTSTSTGPTTVYIGASKTPASIFGPAASHLLILFTLNFTITNLRYEENMPG 73
      159 FTHRSSVSTSTSTGPTTVYIGASKTPASIFGPAASHLLILFTLNFTITNLRYEENMPG 216
Db      74 SRKFNTERLQGLNPLRFNNTSVGPYSCRLTLRPEKDGATGVDALCTHRPDDTGP 133
      219 SRKFNTERLQGLNPLRFNNTSVGPYSCRLTLRPEKDGATGVDALCTHRPDDTGP 278
Qy      134 GLDREQLYLELSQULTHSITELGPTLDRDSL YVNGVFTHRSSVPTSTGVVSEEFITLNT 193
      279 GLDREQLYLELSQULTHSITELGPTLDRDSL YVNGVFTHRSSVPTSTGVVSEEFITLNT 338
Db      194 INNLRVADWGQPGSLKENITDNNMKHLSPLFQSSSLGARVYTCRVIALRSVKNGAETR 253
      339 INNLRVADWGQPGSLKENITDNNMKHLSPLFQSSSLGARVYTCRVIALRSVKNGAETR 398
Qy      254 VDLICTYLOPLSGPGPIKQVFEHLSQQTGTRIGRLGPSIDKDSL YVNGVNEGPDPPT 313
      399 VDLICTYLOPLSGPGPIKQVFEHLSQQTGTRIGRLGPSIDKDSL YVNGVNEGPDPPT 458
Db      314 TPKATTFPLPSEATTAMGYHLKTLTLNFTISNLOYS PDMGKGSATFNSGEGVLQHLR 373
      459 TPKATTFPLPSEATTAMGYHLKTLTLNFTISNLOYS PDMGKGSATFNSGEGVLQHLR 518
Qy      374 PLFQKSMGPFYLGCOLISLRPEKDGATGVDITCTYHPDVGGLDIQQLYWELSQULTH 433
      519 PLFQKSMGPFYLGCOLISLRPEKDGATGVDITCTYHPDVGGLDIQQLYWELSQULTH 578
Db      434 GVTQLGFFVLDROSLFINGYAPQNLSTIRGEYQINFNHVMNLSNPDPSTSSSEYITLLRDIQ 493
      579 GVTQLGFFVLDROSLFINGYAPQNLSTIRGEYQINFNHVMNLSNPDPSTSSSEYITLLRDIQ 638
Qy      494 DKVTLLYKGSQLDHTEFECVLTNLTMDSVLVTVYALFSSNLDPSLVEQVFLDKTLNASFH 553
      639 DKVTLLYKGSQLDHTEFECVLTNLTMDSVLVTVYALFSSNLDPSLVEQVFLDKTLNASFH 698
Db      554 WLGSSTYQLVLDIHVTEMESVYQPTSSSTQHFYINFTITNLPSQDKAQPCTTNYQGNKR 613
      699 WLGSSTYQLVLDIHVTEMESVYQPTSSSTQHFYINFTITNLPSQDKAQPCTTNYQGNKR 758
Qy      614 NIEDALNQLFRNSISIKYFSDCCVSTFRSVNRRHHTGVDSL CNFSPILARRVDAIYEEF 673
      759 NIEDALNQLFRNSISIKYFSDCCVSTFRSVNRRHHTGVDSL CNFSPILARRVDAIYEEF 818
Db      674 LRMTNQTQLQNFLLDRSSVLYDGYSPNREPLTGN 710
      819 LRMTNQTQLQNFLLDRSSVLYDGYSPNREPLTGN 855

```

RESULT 14

ADU01425 ID ADU01425 standard; protein; 914 AA.

AC ADU01425;

DT 30-DEC-2004 (first entry)

XX Breast cancer associated polypeptide seqid 206.

XX Cytostatic; gene therapy; vaccine; breast cancer; cancer; tumour;

XX immune response; OBE; vaccine.

XX Homo sapiens.

PN US2002081609-A1.
 XX 27-JUN-2002.
 PD
 XX 20-JUL-2001; 2001US-00910689.
 PF
 XX 30-NOV-1999; 99US-00451651.
 PR 22-FEB-2000; 2000US-00510662.
 PR 10-MAR-2000; 2000US-00523586.
 PR 07-APR-2000; 2000US-00545068.
 PR 15-MAY-2000; 2000US-00571025.
 PR 06-FEB-2001; 2001US-00778320.
 XX
 PA (DILL/) DILLON D C.
 PA (DAYC/) DAY C H.
 PA (JIAN/) JIANG Y.
 PA (HOUG/) HOUGHTON R L.
 PA (MITC/) MITCHAM J L.
 PA (WANG/) WANG T.
 PA (MCNR/) MCNEILL P D.
 PA (HARL/) HARLOCKER S L.
 PI
 PI Dillion DC, Day CH, Jiang Y, Houghton RL, Mitcham JL, Wang T;
 PI McNeill PD, Harlocker SL;
 XX WPI, 2002-657500/70.
 DR N-PSDB; ADU01424.
 PT Novel breast cancer polynucleotides and polypeptides encoded by the
 PT polynucleotides, useful for detecting the presence of breast cancer in
 PT patient, and in pharmaceutical compositions, for treating breast cancer.
 XX
 PS Claim 2; SEQ ID NO 206; 159pp; English.
 XX
 CC The invention describes an isolated breast cancer polynucleotide (I)
 CC comprising a sequence (S1) selected from the 249 nucleotide sequences
 CC fully defined in the specification, complements of S1, a sequence
 CC comprising at least 20 contiguous residues of S1, a sequence that
 CC hybridize to S1, under moderately stringent conditions, a sequence having
 CC at least 75%, preferably 90% identity to S1, or degenerate variants of
 CC S1. Also described are: an isolated polypeptide (II) encoded by (I), or
 CC sequences having at least 70%, preferably 90% identity to (I); an
 CC expression vector (III); a host cell (IV) transformed or transfected with
 CC (III); an isolated antibody (Ab) that specifically binds to (II);
 CC detecting (M1) the presence of cancer in a patient; a fusion protein (V)
 CC comprising (II); an oligonucleotide (VI) that hybridizes to S1;
 CC stimulating and/or expanding (M2) T cells specific for a tumour protein;
 CC a composition (C) comprising physiologically acceptable carriers and
 CC immunostimulants as first component, and a second component selected from
 CC (I), (II), Ab, (V), (VII) and antigen presenting cells that express (II);
 CC and inhibiting (M3) the development of a cancer in a patient. (C) is
 CC useful for stimulating immune response in a patient, and for treating
 CC cancer in a patient. (VI) is useful for determining the presence of
 CC cancer in a patient, by obtaining a biological sample from the patient,
 CC contacting the biological sample with a monoclonal antibody that bind to
 CC OEB, isolating cells that bind to the antibody that binds to OEB,
 CC isolating polynucleotides from the isolated cells, and contacting the
 CC polynucleotides with (VI), detecting an amount of polynucleotides that
 CC hybridize to the oligonucleotide, and comparing the amount of
 CC polynucleotides that hybridize to the oligonucleotide to a predetermined
 CC cut-off value, and thus determining the presence of cancer in the
 CC patient. (I) and (II) are useful in pharmaceutical compositions, e.g.
 CC vaccines. (I) is useful in the design and preparation of ribozyme
 CC molecules for inhibiting expression of the tumour polypeptides and
 CC proteins in tumour cells. (I) is useful as marker to indicate the
 CC presence or absence of a cancer such as breast cancer. (C) is useful for
 CC inhibiting the development of breast cancer in a patient, and for
 CC removing tumour cells from a biological sample. Ab (binding agent for
 CC (I)) is useful for detecting the presence of cancer in a patient. This
 CC sequence represents a breast cancer associated protein.
 XX
 SEQ Sequence 914 AA;

Query Match 92.9%; Score 3663; DB 5; Length 914;
 Best Local Similarity 99.9%; Pred. No. 2.7e-315;
 Matches 696; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

| | | | |
|----|-----|--|-----|
| QY | 14 | FTHRSSVSTSTPGPTTYLGAASKTPASIFGPSASSHLLIFLTNFTTNLYEENMPG | 73 |
| DB | 159 | FTHRSSVSTSTPGPTTYLGAASKTPASIFGPSASSHLLIFLTNFTTNLYEENMPG | 218 |
| QY | 74 | SRKFNTTTRVYQGLRPLPKQTSVGPLYSGCRLTLRPEKDEATGVDAICTHRDPPTCP | 133 |
| DB | 219 | SRKFNTTTRVYQGLRPLPKQTSVGPLYSGCRLTLRPEKDEATGVDAICTHRDPPTCP | 278 |
| QY | 134 | GLDRPOLYLBLSQTLHSITELGAPYTLDRDSLVLVNGFTHRSSVPTTSTGVSEPTLNT | 193 |
| DB | 279 | GLDRPOLYLBLSQTLHSITELGAPYTLDRDSLVLVNGFTHRSSVPTTSTGVSEPTLNT | 338 |
| QY | 194 | INNLRVYMDMGQPGSLKRNITDNNMKHLLSPFQSSSICGARVTGCRVIALRNVKGAETR | 253 |
| DB | 339 | INNLRVYMDMGQPGSLKRNITDNNMKHLLSPFQSSSICGARVTGCRVIALRNVKGAETR | 398 |
| QY | 254 | VDLLCTYLOPLSGPGLPIKQVFEHSOOTHGITRGPYSLDKDSLVLNGYNEPDPDEPT | 313 |
| DB | 399 | VDLLCTYLOPLSGPGLPIKQVFEHSOOTHGITRGPYSLDKDSLVLNGYNEPDPDEPT | 458 |
| QY | 314 | TPKPATTFPLPLSEATTAMGYHLKTLTNFTNISLQYSPDMGKSGATENSTEGVLOHLIR | 373 |
| DB | 459 | TPKPATTFPLPLSEATTAMGYHLKTLTNFTNISLQYSPDMGKSGATENSTEGVLOHLIR | 518 |
| QY | 374 | PLFQKSSMGPPFLGCOLISLRPEKGAATGVDTCTTHPDPVGPGLDLOQLTWELSQLTH | 433 |
| DB | 519 | PLFQKSSMGPPFLGCOLISLRPEKGAATGVDTCTTHPDPVGPGLDLOQLTWELSQLTH | 578 |
| QY | 434 | GYTOLGFFVLDLDRDSLFINGYAPONLSIRGEYOINFIHVMNLSNPPTSSEYITLLRDIO | 493 |
| DB | 579 | GYTOLGFFVLDLDRDSLFINGYAPONLSIRGEYOINFIHVMNLSNPPTSSEYITLLRDIO | 638 |
| QY | 494 | DKVTLTYKSSQLHDPFRFLCVNTLMDSVLYTVKALFSSNLDPSLVEQVFLDKTLNASH | 553 |
| DB | 639 | DKVTLTYKSSQLHDPFRFLCVNTLMDSVLYTVKALFSSNLDPSLVEQVFLDKTLNASH | 698 |
| QY | 554 | WIGSTYQVLVDIHTEMESSVYOPTSSSSIOHFIYNTINLPISQKAPGTTNYQRNR | 613 |
| DB | 699 | WIGSTYQVLVDIHTEMESSVYOPTSSSSIOHFIYNTINLPISQKAPGTTNYQRNR | 758 |
| QY | 614 | NIEDALNOLFRNSSISYSDCOVSFRSVPNRHHTGVDSLGNFSLARRVDEVALYEEF | 673 |
| DB | 759 | NIEDALNOLFRNSSISYSDCOVSFRSVPNRHHTGVDSLGNFSLARRVDEVALYEEF | 818 |
| QY | 674 | LKMTNRGTQLQNTTLDRSSVLVDGYSPNNEPLTGN | 710 |
| DB | 819 | LKMTNRGTQLQNTTLDRSSVLVDGYSPNNEPLTGN | 855 |

RESULT 15
 ADZ41689
 ID ADZ41689 standard; protein; 914 AA.
 XX
 AC ADZ41689;
 XX
 DT 16-JUN-2005 (first entry)
 XX
 DE Human breast cancer associated protein SEQ ID NO 206.
 XX
 KW cytostatic; gene therapy; vaccine; diagnosis; pharmaceutical; cancer;
 KM neoplasm; breast tumor; endocrine disease; gynecology and obstetrics.
 XX
 OS Homo sapiens.
 XX
 PN W0200262203-A2.
 XX
 PD 15-AUG-2002.
 XX
 PF 04-FEB-2002; 2002MO-US003332.

XX 06-FEB-2001; 2001US-00778320.
 PR 20-JUL-2001; 2001US-00910689.
 PR 30-NOV-2001; 2001US-00010742.
 XX
 XX (CORI-) CORIXA CORP.
 PI Dillion DC, Day CH, Jiang Y, Houghton RL, Mitcham JL, Wang T;
 PI Menelli PD, Harlocker SL, Bennington AA, Zehentner B, Fanger GR;
 PI Retter MW;
 XX
 DR MPI; 2002-657500/70.
 DR N-PSDB; ADZ41688.
 XX
 PT Novel breast cancer polynucleotides and polypeptides encoded by the
 PT polynucleotides, useful for detecting the presence of breast cancer in
 PT patient, and in pharmaceutical compositions, for treating breast cancer.
 XX
 PS Claim 2; SEQ ID NO 206, 237pp; English.
 XX
 CC The invention describes an isolated breast cancer polynucleotide (I)
 CC comprising a sequence (S1) selected from the 249 nucleotide sequences
 CC fully defined in the specification, complements of S1, a sequence
 CC comprising at least 20 contiguous residues of S1, a sequence that
 CC hybridize to S1, under moderately stringent conditions, a sequence having
 CC at least 75%, preferably 90% identity to S1, or degenerate variants of
 CC S1. Also described are: an isolated polypeptide (II) comprising a
 CC sequence (S2) selected from any one of the 11 sequences mentioned in the
 CC specification, sequences encoded by (I), or sequences having at least
 CC 70%, preferably 90% identity to (I); an expression vector (III)
 CC comprising (I), operably linked to an expression control sequence; a host
 CC cell (IV) transformed or transfected with (III); an isolated antibody
 CC (Ab) or its antigen-binding fragment, that specifically binds to (II);
 CC detecting (M1) the presence of cancer in a patient; a fusion protein (V)
 CC comprising (II); an oligonucleotide (VI) that hybridizes to S1;
 CC stimulating and/or expanding (M2) T cells specific for a tumor protein;
 CC an isolated T cell population (VII), comprising T cells prepared by (M2);
 CC a composition (C) comprising physiologically acceptable carriers and
 CC immunostimulants as first component, and a second component selected from
 CC (I), (II), Ab, (V), (VII) and antigen presenting cells that express (II);
 CC a diagnostic kit (VIII) comprising (VI), or Ab and a detection reagent
 CC comprising a reporter group; and inhibiting (M3) the development of a
 CC cancer in a patient, by incubating CD4+ and/or CD8+ T cells isolated from
 CC a patient with (I), (II) or antigen presenting cells that express (II),
 CC so that T cell proliferate, administering the proliferated T cells to the
 CC patient, and therefore inhibiting the development of a cancer in the
 CC patient. The following are disclosed: monitoring the progression of
 CC cancer; fragments of (II); (xenogeneic) variants of (I); polynucleotide
 CC compositions comprising antisense oligonucleotide; and kit for use in
 CC diagnostic methods. (C) is useful for stimulating immune response in a
 CC patient, and for treating cancer in a patient. (VI) is useful for
 CC determining the presence of cancer in a patient. (I) and (II) are useful
 CC in pharmaceutical compositions, e.g. vaccines. (I) is useful in the
 CC design and preparation of ribozyme molecules for inhibiting expression of
 CC the tumor polypeptides and proteins in tumor cells. (I) is useful as
 CC marker to indicate the presence or absence of a cancer such as breast
 CC cancer. (C) is useful for inhibiting the development of breast cancer in
 CC a patient, and for removing tumor cells from a biological sample. Ab
 CC (binding agent for (I)) is useful for detecting the presence of cancer in
 CC a patient. This is the amino acid sequence of a breast cancer associated
 CC protein. Note: This sequence has been extracted from the sequence listing
 CC of a corrected version of the specification published on the 3rd of March
 CC 2005.
 XX
 SO Sequence 914 AA:

Query Match 92.9%; Score 3663; DB 5; Length 914;
 Best Local Similarity 99.9%; Pred. No. 2.7e-315;
 Matches 696; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

14 FTNRSSVSTSTGPTVYVYAGASKTPASIRGSAASHLLFTLNFTITNLRKYENNMWPG 73
 |||||||
 DB 159 FTHRSSVSTSTGPTVYVYAGASKTPASIRGSAASHLLFTLNFTITNLRKYENNMWPG 218

QY 74 SRKNTTERVLOGLRLPLFKNTSVGLYSGCRLLTLRPEKDGATGVDAICTHRPDPYGP 133
 |||||||
 DB 219 SRKNTTERVLOGLRLPLFKNTSVGLYSGCRLLTLRPEKDGATGVDAICTHRPDPYGP 278
 |||||||
 QY 134 GLDRBQLYELSQLTHSITELGPTLDRSLVYNGFTHRSSVPTSTGVSEEPFTLNFT 193
 |||||||
 DB 279 GLDRBQLYELSQLTHSITELGPTLDRSLVYNGFTHRSSVPTSTGVSEEPFTLNFT 338
 |||||||
 QY 194 INNLRYMADMGQPGSLKFNITDVMKHLISPLFQSSSLGARVYGCYVALRSYKGAETR 253
 |||||||
 DB 339 INNLRYMADMGQPGSLKFNITDVMKHLISPLFQSSSLGARVYGCYVALRSYKGAETR 398
 |||||||
 QY 254 VDLCTTYLOPLSGRGLPIKQVPHLSQQTHGTRGLGPLYLDKDSLVLNGNEGDPPEPT 313
 |||||||
 DB 399 VDLCTTYLOPLSGRGLPIKQVPHLSQQTHGTRGLGPLYLDKDSLVLNGNEGDPPEPT 458
 |||||||
 QY 314 TPKPATTPPLSEATTAMGYHLKTLTLNFTSNLQSPDMKGSATFNSBEGVLOHLRL 373
 |||||||
 DB 459 TPKPATTPPLSEATTAMGYHLKTLTLNFTSNLQSPDMKGSATFNSBEGVLOHLRL 518
 |||||||
 QY 374 PLFQKSSMGPFYLGCOLISLRPEKDGATGVDTCTYHPDPVPGGLDIQOLYWELSQLT 433
 |||||||
 DB 519 PLFQKSSMGPFYLGCOLISLRPEKDGATGVDTCTYHPDPVPGGLDIQOLYWELSQLT 578
 |||||||
 QY 434 GYTQLGPLYLDRDSLFINGYAPQNLISRGYQINFIYVMNLSNPDPTSEYITLLRDIO 493
 |||||||
 DB 579 GYTQLGPLYLDRDSLFINGYAPQNLISRGYQINFIYVMNLSNPDPTSEYITLLRDIO 638
 |||||||
 QY 494 DKYTLTKGSQLHDFRFLVTLNLTMDSVLNTVKALFSSNLDPDLVEQVFLDKTLNLSF 553
 |||||||
 DB 639 DKYTLTKGSQLHDFRFLVTLNLTMDSVLNTVKALFSSNLDPDLVEQVFLDKTLNLSF 698
 |||||||
 QY 554 WLGSTYQVLDIHVTMESSVYOPTSSSTQHFYLNFTINLPSYQDKAPGTINYORNR 613
 |||||||
 DB 699 WLGSTYQVLDIHVTMESSVYOPTSSSTQHFYLNFTINLPSYQDKAPGTINYORNR 758
 |||||||
 QY 614 NIEDALNQLFRSSISYSDCOVSTFRSVPNRHHTGVSLCNFSLARVDVAIYEER 673
 |||||||
 DB 759 NIEDALNQLFRSSISYSDCOVSTFRSVPNRHHTGVSLCNFSLARVDVAIYEER 818
 |||||||
 QY 674 LMRTRNGTOLQNPFTLDRSSVLVDGYSPPNRNEPLTGN 710
 |||||||
 DB 819 LMRTRNGTOLQNPFTLDRSSVLVDGYSPPNRNEPLTGN 855
 |||||||

Search completed: October 14, 2006, 03:59:00
 Job time : 208 secs

THIS PAGE BLANK (USPTO)

QY 301 NCYNBGPDEPPTPKPATFLPLPSEATTAMGYHLKTLTLPNTISNLQYSPDMGKSAT 360
DB 301 NCYNBGPDEPPTPKPATFLPLPSEATTAMGYHLKTLTLPNTISNLQYSPDMGKSAT 360
QY 361 FSTSGVLOHLRLPFLQKSMGPFYVGCCLISLRPEKDAATGVDTTCYHDPVPGD 420
DB 361 FSTSGVLOHLRLPFLQKSMGPFYVGCCLISLRPEKDAATGVDTTCYHDPVPGD 420
QY 421 IOOLYWEISQLTGVTQGLFVYLDLDRSLFINGYAPONLSIRGEYOQINFIHVMNLSNPD 480
DB 421 IOOLYWEISQLTGVTQGLFVYLDLDRSLFINGYAPONLSIRGEYOQINFIHVMNLSNPD 480
QY 481 TSSEYITLLRDIQDKVTTLYKGSQQLHDTFRFCLVTNLTMDSVLYVKALFSSNLDPSLVE 540
DB 481 TSSEYITLLRDIQDKVTTLYKGSQQLHDTFRFCLVTNLTMDSVLYVKALFSSNLDPSLVE 540
QY 541 QVFLDKTLNASFHMLGASTYQVLDIHVTEMESSVYQPTSSSTQHFLYNTITNLPSYQDK 600
DB 541 QVFLDKTLNASFHMLGASTYQVLDIHVTEMESSVYQPTSSSTQHFLYNTITNLPSYQDK 600
QY 601 AOPGTNYQRNKNIEDALNQLFRNSSIKSYFSDCOVSTFRSVPNRHHTGVDSLGNFSP 660
DB 601 AOPGTNYQRNKNIEDALNQLFRNSSIKSYFSDCOVSTFRSVPNRHHTGVDSLGNFSP 660
QY 661 ARVDVAIYEEFLRMTNRGTQLQNTLDRSSVLYVGYSPNREPLTGNSADIQHSQGRS 720
DB 661 ARVDVAIYEEFLRMTNRGTQLQNTLDRSSVLYVGYSPNREPLTGNSADIQHSQGRS 720
QY 721 SLEGPFEOKLISEEDLNMTGHHHHH 748
DB 721 SLEGPFEOKLISEEDLNMTGHHHHH 748

RESULT 2
US-10-687-035-2
Sequence 2, Application US/10687035
Publication No. US20050064518A1
GENERAL INFORMATION:
APPLICANT: Albione, Earl F.
APPLICANT: Solitis, Daniel A.
TITLE OF INVENTION: ANTIBODIES THAT BIND CELL-ASSOCIATED
FILE REFERENCE: 6750-214-999
CURRENT APPLICATION NUMBER: US/10/687,035
CURRENT FILING DATE: 2003-10-15
PRIOR APPLICATION NUMBER: 60/485,986
PRIOR FILING DATE: 2003-07-10
PRIOR APPLICATION NUMBER: 60/418,828
PRIOR FILING DATE: 2003-10-12
NUMBER OF SEQ ID NOS: 71
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 809
TYPE: PRN
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: CA 125/0772P 3-repeat TM
US-10-687-035-2

Query Match 99.0%; Score 3904.5; DB 5; Length 809;
Best Local Similarity 92.5%; Pred. No. 7.7e-309;
Matches 748; Conservative 0; Mismatches 0; Indels 61; Gaps 1;
QY 1 AAOBARARRTKLTTHSSVSTSTPCTPTVYLGASCTPASPISFQSAASHLLILFTLNT 60
DB 1 AAOBARARRTKLTTHSSVSTSTPCTPTVYLGASCTPASPISFQSAASHLLILFTLNT 60
QY 61 ITNLRYEENMPPGSRKNTTERVLOGLRLPFXNTSVGPVLYSGCRLTLRPEKQGEATGV 120
DB 61 ITNLRYEENMPPGSRKNTTERVLOGLRLPFXNTSVGPVLYSGCRLTLRPEKQGEATGV 120
QY 121 DATCTHAPDPTGRLDREQLYLELSQLTGATLGRDLSLYVNGFTHRSVPPTST 180
DB 121 DATCTHAPDPTGRLDREQLYLELSQLTGATLGRDLSLYVNGFTHRSVPPTST 180

DB 121 DATCTHAPDPTGRLDREQLYLELSQLTGATLGRDLSLYVNGFTHRSVPPTST 180
QY 181 GVYSEEPPLTANTTNKLYAMDMGQPSLKFNITNNMKHLLSPFORSSILARYTGCV 240
DB 181 GVYSEEPPLTANTTNKLYAMDMGQPSLKFNITNNMKHLLSPFORSSILARYTGCV 240
QY 241 IALRSYKNGAETRVLLCTYLOPLSGPGLPIQVFEHLSQOHTGTRLGPSLDKXSLYL 300
DB 241 IALRSYKNGAETRVLLCTYLOPLSGPGLPIQVFEHLSQOHTGTRLGPSLDKXSLYL 300
QY 301 NCYNBGPDEPPTPKPATFLPLPSEATTAMGYHLKTLTLPNTISNLQYSPDMGKSAT 360
DB 301 NCYNBGPDEPPTPKPATFLPLPSEATTAMGYHLKTLTLPNTISNLQYSPDMGKSAT 360
QY 361 FSTSGVLOHLRLPFLQKSMGPFYVGCCLISLRPEKDAATGVDTTCYHDPVPGD 420
DB 361 FSTSGVLOHLRLPFLQKSMGPFYVGCCLISLRPEKDAATGVDTTCYHDPVPGD 420
QY 421 IOOLYWEISQLTGVTQGLFVYLDLDRSLFINGYAPONLSIRGEYOQINFIHVMNLSNPD 480
DB 421 IOOLYWEISQLTGVTQGLFVYLDLDRSLFINGYAPONLSIRGEYOQINFIHVMNLSNPD 480
QY 481 TSSEYITLLRDIQDKVTTLYKGSQQLHDTFRFCLVTNLTMDSVLYVKALFSSNLDPSLVE 540
DB 481 TSSEYITLLRDIQDKVTTLYKGSQQLHDTFRFCLVTNLTMDSVLYVKALFSSNLDPSLVE 540
QY 541 QVFLDKTLNASFHMLGASTYQVLDIHVTEMESSVYQPTSSSTQHFLYNTITNLPSYQDK 600
DB 541 QVFLDKTLNASFHMLGASTYQVLDIHVTEMESSVYQPTSSSTQHFLYNTITNLPSYQDK 600
QY 601 AOPGTNYQRNKNIEDALNQLFRNSSIKSYFSDCOVSTFRSVPNRHHTGVDSLGNFSP 660
DB 601 AOPGTNYQRNKNIEDALNQLFRNSSIKSYFSDCOVSTFRSVPNRHHTGVDSLGNFSP 660
QY 661 ARVDVAIYEEFLRMTNRGTQLQNTLDRSSVLYVGYSPNREPLTGNSADIQHSQGRS 720
DB 661 ARVDVAIYEEFLRMTNRGTQLQNTLDRSSVLYVGYSPNREPLTGNSADIQHSQGRS 720
QY 709 -----NSADIQHSQGR 719
DB 721 GLAGLGLITCLIGVLTTRRRKGEYVNOQCPGYOSHLDLDELQNSADIQHSQGR 780
QY 720 SLEGPFEOKLISEEDLNMTGHHHHH 748
DB 781 SLEGPFEOKLISEEDLNMTGHHHHH 809

RESULT 3
US-10-983-340-4
Sequence 4, Application US/10983340
Publication No. US20050238649A1
GENERAL INFORMATION:
APPLICANT: Doronina, Svetlana O.
APPLICANT: Senter, Peter D.
APPLICANT: Ebens, Allen J.
APPLICANT: Polakis, Paul
APPLICANT: Sliwkowski, Mark X.
APPLICANT: Spencer, Susan D.
APPLICANT: Kline, Toni Beth
TITLE OF INVENTION: MONOMETHYLAVALINE COMPOUNDS CAPABLE OF CONJUGATION TO LIGANDS
FILE REFERENCE: 018691-001020US
CURRENT APPLICATION NUMBER: US/10/983,340
CURRENT FILING DATE: 2004-11-05
PRIOR APPLICATION NUMBER: US 60/598,899
PRIOR FILING DATE: 2004-08-04
PRIOR APPLICATION NUMBER: US 60/557,116
PRIOR FILING DATE: 2004-03-26
PRIOR APPLICATION NUMBER: US 60/518,534
PRIOR FILING DATE: 2003-11-06
NUMBER OF SEQ ID NOS: 35
SEQ ID NO 4
LENGTH: 6995

TYPE: PRT
ORGANISM: Homo sapien
US-10-983-340-4

Query Match 92.9%; Score 3665; DB 5; Length 6995;
Best Local Similarity 99.9%; Pred. No. 7.5e-288;
Matches 696; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
QY 14 FTHRSSVSTSTPGTPVYVIGASKTPASIFGSPAASHLLILFTLNFTITNLRYEENMPG 73
DB 6240 FTHRSSVSTSTPGTPVYVIGASKTPASIFGSPAASHLLILFTLNFTITNLRYEENMPG 6299
QY 74 SRKFNTERVLOGLRLPLFKNTSVGPYSGCRITLLRPEKDGATGVDALCTHRPDPGP 133
DB 6300 SRKFNTERVLOGLRLPLFKNTSVGPYSGCRITLLRPEKDGATGVDALCTHRPDPGP 6359
QY 134 GLDREQLYLELSQUTHSITELGPTYLDRDSLTVNGFTHRSSVPTSTGVVSEBFTLNFT 193
DB 6360 GLDREQLYLELSQUTHSITELGPTYLDRDSLTVNGFTHRSSVPTSTGVVSEBFTLNFT 6419
QY 194 INNLRYADMGQPSGLKFNITTDVNMKHLSPFORSSLGARYTCRVIALRSVNGAETR 253
DB 6420 INNLRYADMGQPSGLKFNITTDVNMKHLSPFORSSLGARYTCRVIALRSVNGAETR 6479
QY 254 VDLCTYLOPLSGRGLPIKOVFHELSQOOTHGITRLGPFYSLDKDSLTVNGVNEBGPDPPT 313
DB 6480 VDLCTYLOPLSGRGLPIKOVFHELSQOOTHGITRLGPFYSLDKDSLTVNGVNEBGPDPPT 6539
QY 314 TPXPATTFPLPLSEATTAMGYHLKTLTNFTISNLQYSPDMGKGSATFNSTEGVLOHLR 373
DB 6540 TPXPATTFPLPLSEATTAMGYHLKTLTNFTISNLQYSPDMGKGSATFNSTEGVLOHLR 6599
QY 374 PLFOKSMGPFYVGCQQLISLRPEKDGATGVDTCYHPDPVPGGLDIQQLYHELSQLTH 433
DB 6600 PLFOKSMGPFYVGCQQLISLRPEKDGATGVDTCYHPDPVPGGLDIQQLYHELSQLTH 6659
QY 434 GVTQLGFFVYLDRLSLFNGYAPQNLSTRGEYQINPHVNMNLSNPDPSTSEYITLLDIO 493
DB 6660 GVTQLGFFVYLDRLSLFNGYAPQNLSTRGEYQINPHVNMNLSNPDPSTSEYITLLDIO 6719
QY 494 DKYTTLYKGSQLDHTRFCLVTNLTMDSVLYTVKALFSSNLDPSELVEQVFLDKTLNLSFH 553
DB 6720 DKYTTLYKGSQLDHTRFCLVTNLTMDSVLYTVKALFSSNLDPSELVEQVFLDKTLNLSFH 6779
QY 554 WLGSTYQVLDIHVTEMESYVQPTSSSTQHFYLNFTITNLPSQDXYAQPGTNYQNK 613
DB 6780 WLGSTYQVLDIHVTEMESYVQPTSSSTQHFYLNFTITNLPSQDXYAQPGTNYQNK 6839
QY 614 NIDDALNQLFRNSISIKYFSDCOVSTFRSVPNRHHTGVDSL CNFSPILARRVDAIYE 673
DB 6840 NIDDALNQLFRNSISIKYFSDCOVSTFRSVPNRHHTGVDSL CNFSPILARRVDAIYE 6899
QY 674 LRMTNRGTQLONFTLDRSSVLYVDGYSNNRNEPLTGN 710
DB 6900 LRMTNRGTQLONFTLDRSSVLYVDGYSNNRNEPLTGN 6936
```

RESULT 4
US-09-884-441-389
Sequence 389, Application US/09884441
Patent No. US20020119158A1
GENERAL INFORMATION:
APPLICANT: Algate, Paul A.
APPLICANT: Carter, Darrick
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.462C7
CURRENT APPLICATION NUMBER: US/09/884,441
CURRENT FILING DATE: 2001-06-18
NUMBER OF SEQ ID NOS: 489
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 389
LENGTH: 833

TYPE: PRT
ORGANISM: Homo sapiens
US-09-884-441-389

Query Match 92.9%; Score 3663; DB 3; Length 833;
Best Local Similarity 99.9%; Pred. No. 4e-289;
Matches 696; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 14 FTHRSSVSTSTPGTPVYVIGASKTPASIFGSPAASHLLILFTLNFTITNLRYEENMPG 73
DB 78 FTHRSSVSTSTPGTPVYVIGASKTPASIFGSPAASHLLILFTLNFTITNLRYEENMPG 137
QY 74 SRKFNTERVLOGLRLPLFKNTSVGPYSGCRITLLRPEKDGATGVDALCTHRPDPGP 133
DB 138 SRKFNTERVLOGLRLPLFKNTSVGPYSGCRITLLRPEKDGATGVDALCTHRPDPGP 197
QY 134 GLDREQLYLELSQUTHSITELGPTYLDRDSLTVNGFTHRSSVPTSTGVVSEBFTLNFT 193
DB 198 GLDREQLYLELSQUTHSITELGPTYLDRDSLTVNGFTHRSSVPTSTGVVSEBFTLNFT 257
QY 194 INNLRYADMGQPSGLKFNITTDVNMKHLSPFORSSLGARYTCRVIALRSVNGAETR 253
DB 258 INNLRYADMGQPSGLKFNITTDVNMKHLSPFORSSLGARYTCRVIALRSVNGAETR 317
QY 254 VDLCTYLOPLSGRGLPIKOVFHELSQOOTHGITRLGPFYSLDKDSLTVNGVNEBGPDPPT 313
DB 318 VDLCTYLOPLSGRGLPIKOVFHELSQOOTHGITRLGPFYSLDKDSLTVNGVNEBGPDPPT 377
QY 314 TPXPATTFPLPLSEATTAMGYHLKTLTNFTISNLQYSPDMGKGSATFNSTEGVLOHLR 373
DB 378 TPXPATTFPLPLSEATTAMGYHLKTLTNFTISNLQYSPDMGKGSATFNSTEGVLOHLR 437
QY 374 PLFOKSMGPFYVGCQQLISLRPEKDGATGVDTCYHPDPVPGGLDIQQLYHELSQLTH 433
DB 438 PLFOKSMGPFYVGCQQLISLRPEKDGATGVDTCYHPDPVPGGLDIQQLYHELSQLTH 497
QY 434 GVTQLGFFVYLDRLSLFNGYAPQNLSTRGEYQINPHVNMNLSNPDPSTSEYITLLDIO 493
DB 498 GVTQLGFFVYLDRLSLFNGYAPQNLSTRGEYQINPHVNMNLSNPDPSTSEYITLLDIO 557
QY 494 DKYTTLYKGSQLDHTRFCLVTNLTMDSVLYTVKALFSSNLDPSELVEQVFLDKTLNLSFH 553
DB 558 DKYTTLYKGSQLDHTRFCLVTNLTMDSVLYTVKALFSSNLDPSELVEQVFLDKTLNLSFH 617
QY 554 WLGSTYQVLDIHVTEMESYVQPTSSSTQHFYLNFTITNLPSQDXYAQPGTNYQNK 613
DB 618 WLGSTYQVLDIHVTEMESYVQPTSSSTQHFYLNFTITNLPSQDXYAQPGTNYQNK 677
QY 614 NIDDALNQLFRNSISIKYFSDCOVSTFRSVPNRHHTGVDSL CNFSPILARRVDAIYE 673
DB 678 NIDDALNQLFRNSISIKYFSDCOVSTFRSVPNRHHTGVDSL CNFSPILARRVDAIYE 737
QY 674 LRMTNRGTQLONFTLDRSSVLYVDGYSNNRNEPLTGN 710
DB 738 LRMTNRGTQLONFTLDRSSVLYVDGYSNNRNEPLTGN 774
```

RESULT 5
US-09-907-969-389
Sequence 389, Application US/09907969
Publication No. US20030091580A1
GENERAL INFORMATION:
APPLICANT: Mitcham, Jennifer L.
APPLICANT: King, Gordon E.
APPLICANT: Algate, Paul A.
APPLICANT: Piling, Steven P.
APPLICANT: Reiter, Marc W.
APPLICANT: Ranger, Gary Richard
APPLICANT: Reed, Steven G.
APPLICANT: Vedvick, Thomas S.
APPLICANT: Carter, Darrick
APPLICANT: Hill, Paul
APPLICANT: Albane, Earl

;; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
;; FILE REFERENCE: 210121.462C6
;; CURRENT APPLICATION NUMBER: US/09/907,969
;; CURRENT FILING DATE: 2001-07-17
;; NUMBER OF SEQ ID NOS: 596
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 389
;; LENGTH: 833
;; TYPE: PR1
;; ORGANISM: Homo sapiens
US-09-907-969-389

Query Match 92.9%; Score 3663; DB 3; Length 833;
Best Local Similarity 99.9%; Pred. No. 4e-289;
Matches 696; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 14 FTHRSVSTSTPGTPTVYLGASKTSPASIFGSAASHLLILFTLNFTTNLRYEENMMPG 73
DB 78 FTHRSVSTSTPGTPTVYLGASKTSPASIFGSAASHLLILFTLNFTTNLRYEENMMPG 137

QY 74 SRKFNTTERVLOGLRPLFKNTSVGPLYSGCRLLTLRPEKDGATGVDAICTHRPDPTGP 133
DB 138 SRKFNTTERVLOGLRPLFKNTSVGPLYSGCRLLTLRPEKDGATGVDAICTHRPDPTGP 197

QY 134 GLDREOLYLESQLTHSITELGPTLDRDSLVLNGFTHRSVPTSTGVVSEEPFLNTFT 193
DB 198 GLDREOLYLESQLTHSITELGPTLDRDSLVLNGFTHRSVPTSTGVVSEEPFLNTFT 257

QY 194 INNLRYMADMGQPGSLKFNITDNVMKHLISPLFORSSIGARYTGCRVIALRSYONCAETR 253
DB 258 INNLRYMADMGQPGSLKFNITDNVMKHLISPLFORSSIGARYTGCRVIALRSYONCAETR 317

QY 254 VDLICTYLOPLSGPGLPIKQVFHLSQOHTGTRIGLAPYSLDKDSLVLNGNEBGPDEPPT 313
DB 318 VDLICTYLOPLSGPGLPIKQVFHLSQOHTGTRIGLAPYSLDKDSLVLNGNEBGPDEPPT 377

QY 314 TPKPATTFPLPSEATTAMGYHLKTLTNFTISNLQYSPDMGKGSATFNSTBGLQHLR 373
DB 378 TPKPATTFPLPSEATTAMGYHLKTLTNFTISNLQYSPDMGKGSATFNSTBGLQHLR 437

QY 374 PLFOKSMGPFYLGCOLISLRPEKGAATGVDITCTYHPDPVPGDLIOQLYELSQLTH 433
DB 438 PLFOKSMGPFYLGCOLISLRPEKGAATGVDITCTYHPDPVPGDLIOQLYELSQLTH 497

QY 434 GTVQLGFFYVLDRLSLFINGYAPONLSIRGEYOINFHIVNMNLSNPPTSSEYITLLRDIO 493
DB 498 GTVQLGFFYVLDRLSLFINGYAPONLSIRGEYOINFHIVNMNLSNPPTSSEYITLLRDIO 557

QY 494 DKYTTILYKGSQQLHDTFRFCVLTNLIMDSVLYTVKALFSSNLDPSLVEQVFLDKTLNASFH 553
DB 558 DKYTTILYKGSQQLHDTFRFCVLTNLIMDSVLYTVKALFSSNLDPSLVEQVFLDKTLNASFH 617

QY 554 WLGSSTYQLVDIHTEMSSVYOPTSSSTQHFYLNFTTNLPYSQDAQAGTTNYQNK 613
DB 618 WLGSSTYQLVDIHTEMSSVYOPTSSSTQHFYLNFTTNLPYSQDAQAGTTNYQNK 677

QY 614 NIEDALNQLFRNSISIKSYFSDCOVSTFRSVPNRHHTGVDSL CNFSPILARRVDVAIYEEF 673
DB 678 NIEDALNQLFRNSISIKSYFSDCOVSTFRSVPNRHHTGVDSL CNFSPILARRVDVAIYEEF 737

QY 674 LRMTNRGTOLONFTLDRSSVLYVDGYFPNRNEPPLTGN 710
DB 738 LRMTNRGTOLONFTLDRSSVLYVDGYFPNRNEPPLTGN 774

RESULT 6
US-09-827-271-389
; Sequence 389, Application US/0982721
; Publication No. US20030165504A1
; GENERAL INFORMATION:
; APPLICANT: Retter, Marc W.
; APPLICANT: Fanger, Gary R.

;; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;; TITLE OF INVENTION: DIAGNOSIS OF OVARIAN CANCER
;; FILE REFERENCE: 210121.462C6
;; CURRENT APPLICATION NUMBER: US/09/827,271
;; CURRENT FILING DATE: 2001-04-04
;; NUMBER OF SEQ ID NOS: 461
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 389
;; LENGTH: 833
;; TYPE: PR1
;; ORGANISM: Homo sapiens
US-09-827-271-389

Query Match 92.9%; Score 3663; DB 3; Length 833;
Best Local Similarity 99.9%; Pred. No. 4e-289;
Matches 696; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 14 FTHRSVSTSTPGTPTVYLGASKTSPASIFGSAASHLLILFTLNFTTNLRYEENMMPG 73
DB 78 FTHRSVSTSTPGTPTVYLGASKTSPASIFGSAASHLLILFTLNFTTNLRYEENMMPG 137

QY 74 SRKFNTTERVLOGLRPLFKNTSVGPLYSGCRLLTLRPEKDGATGVDAICTHRPDPTGP 133
DB 138 SRKFNTTERVLOGLRPLFKNTSVGPLYSGCRLLTLRPEKDGATGVDAICTHRPDPTGP 197

QY 134 GLDREOLYLESQLTHSITELGPTLDRDSLVLNGFTHRSVPTSTGVVSEEPFLNTFT 193
DB 198 GLDREOLYLESQLTHSITELGPTLDRDSLVLNGFTHRSVPTSTGVVSEEPFLNTFT 257

QY 194 INNLRYMADMGQPGSLKFNITDNVMKHLISPLFORSSIGARYTGCRVIALRSYONCAETR 253
DB 258 INNLRYMADMGQPGSLKFNITDNVMKHLISPLFORSSIGARYTGCRVIALRSYONCAETR 317

QY 254 VDLICTYLOPLSGPGLPIKQVFHLSQOHTGTRIGLAPYSLDKDSLVLNGNEBGPDEPPT 313
DB 318 VDLICTYLOPLSGPGLPIKQVFHLSQOHTGTRIGLAPYSLDKDSLVLNGNEBGPDEPPT 377

QY 314 TPKPATTFPLPSEATTAMGYHLKTLTNFTISNLQYSPDMGKGSATFNSTBGLQHLR 373
DB 378 TPKPATTFPLPSEATTAMGYHLKTLTNFTISNLQYSPDMGKGSATFNSTBGLQHLR 437

QY 374 PLFOKSMGPFYLGCOLISLRPEKGAATGVDITCTYHPDPVPGDLIOQLYELSQLTH 433
DB 438 PLFOKSMGPFYLGCOLISLRPEKGAATGVDITCTYHPDPVPGDLIOQLYELSQLTH 497

QY 434 GTVQLGFFYVLDRLSLFINGYAPONLSIRGEYOINFHIVNMNLSNPPTSSEYITLLRDIO 493
DB 498 GTVQLGFFYVLDRLSLFINGYAPONLSIRGEYOINFHIVNMNLSNPPTSSEYITLLRDIO 557

QY 494 DKYTTILYKGSQQLHDTFRFCVLTNLIMDSVLYTVKALFSSNLDPSLVEQVFLDKTLNASFH 553
DB 558 DKYTTILYKGSQQLHDTFRFCVLTNLIMDSVLYTVKALFSSNLDPSLVEQVFLDKTLNASFH 617

QY 554 WLGSSTYQLVDIHTEMSSVYOPTSSSTQHFYLNFTTNLPYSQDAQAGTTNYQNK 613
DB 618 WLGSSTYQLVDIHTEMSSVYOPTSSSTQHFYLNFTTNLPYSQDAQAGTTNYQNK 677

QY 614 NIEDALNQLFRNSISIKSYFSDCOVSTFRSVPNRHHTGVDSL CNFSPILARRVDVAIYEEF 673
DB 678 NIEDALNQLFRNSISIKSYFSDCOVSTFRSVPNRHHTGVDSL CNFSPILARRVDVAIYEEF 737

QY 674 LRMTNRGTOLONFTLDRSSVLYVDGYFPNRNEPPLTGN 710
DB 738 LRMTNRGTOLONFTLDRSSVLYVDGYFPNRNEPPLTGN 774

RESULT 7
US-10-198-053-389
; Sequence 389, Application US/10198053
; Publication No. US20030124140A1
; GENERAL INFORMATION:
; APPLICANT: Bangur, Chalcanya S.
; APPLICANT: Retter, Marc W.

APPLICANT: Fanger, Gary R.
APPLICANT: Hill, Paul
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
FILE REFERENCE: 210121.462C9
CURRENT APPLICATION NUMBER: US/10/198,053
CURRENT FILING DATE: 2002-07-17
NUMBER OF SEQ ID NOS: 624
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 389
LENGTH: 833
TYPE: PRT
ORGANISM: Homo sapiens
US-10-198-053-389

Query Match 92.9%; Score 3663; DB 4; Length 833;
Best Local Similarity 99.9%; Pred. No. 4e-289;
Matches 696; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

14 FTHRSSVSTSTGTPTVYIGASKTPASIFGPSASHLLILFTLNFTTNLRYEENWPG 73
78 FTHRSSVSTSTGTPTVYIGASKTPASIFGPSASHLLILFTLNFTTNLRYEENWPG 137
74 SRKFNTERVLOGLRLPLFKNTSVGPLYSGCRLTLRPEKDEATGVDAICTHRPDTGP 133
138 SRKFNTERVLOGLRLPLFKNTSVGPLYSGCRLTLRPEKDEATGVDAICTHRPDTGP 197
134 GDREOLYLELSQTLHSITELGPTTLDRLSYNGFTHRSSVPTTSGVSEEPFTLNFT 193
198 GDREOLYLELSQTLHSITELGPTTLDRLSYNGFTHRSSVPTTSGVSEEPFTLNFT 257
194 INNLRYADMGQSGSLKFNITDVMKHLSPFORSSLGARYTCRVALRSYKNGAETR 253
258 INNLRYADMGQSGSLKFNITDVMKHLSPFORSSLGARYTCRVALRSYKNGAETR 317
254 VDLCTYLOPLSGRGLPIKQVFEHLSQOHTGIRLGRPLSDKSLYNGNEGPDPPT 313
318 VDLCTYLOPLSGRGLPIKQVFEHLSQOHTGIRLGRPLSDKSLYNGNEGPDPPT 377
314 TPXPATTFPLPSEATTAMGYHLKTLTNFTISNLQYSPDMGKSATFNSTEGVLOHLR 373
378 TPXPATTFPLPSEATTAMGYHLKTLTNFTISNLQYSPDMGKSATFNSTEGVLOHLR 437
374 PLFOKSMGPFYIGCOLISLRPEKGAATGVDTCTYHDPVPGGLDIQOLYWELSQTLH 433
438 PLFOKSMGPFYIGCOLISLRPEKGAATGVDTCTYHDPVPGGLDIQOLYWELSQTLH 497
434 GVTOLGFYVLDRLSLFINGAPONLSIRGEYQINFIHVMNLSNPDPSTSEYITLLRDIO 493
498 GVTOLGFYVLDRLSLFINGAPONLSIRGEYQINFIHVMNLSNPDPSTSEYITLLRDIO 557
494 DKYTTLYKSQLDHTFRFCVLTNLTMDSVLVTKALFSSNLDPDLVEOVFLDXTLNASFH 553
558 DKYTTLYKSQLDHTFRFCVLTNLTMDSVLVTKALFSSNLDPDLVEOVFLDXTLNASFH 617
554 WLGSTYQVLDIVHTEMESSYVOPTSSSTQHFYLNFTTNLPYSQDKAOPGTTNYQRNKR 613
618 WLGSTYQVLDIVHTEMESSYVOPTSSSTQHFYLNFTTNLPYSQDKAOPGTTNYQRNKR 677
614 NIEDALNOLFRRNSISYSDCOVSTFRSVPNRHHGTGVDSLCNFSLARRVDRVAIYEFP 673
678 NIEDALNOLFRRNSISYSDCOVSTFRSVPNRHHGTGVDSLCNFSLARRVDRVAIYEFP 737
674 LRMTNRGTOLQNFLLDRSSVLDVGYSPNREPLTGN 710
738 LRMTNRGTOLQNFLLDRSSVLDVGYSPNREPLTGN 774

RESULT 8
US-10-860-790-389
Sequence 389, Application US/10860790
Publication No. US20050031634A1
GENERAL INFORMATION:

APPLICANT: Bangur, Chalcanya S.
APPLICANT: Retter, Marc W.
APPLICANT: Fanger, Gary R.
APPLICANT: Hill, Paul
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
FILE REFERENCE: 210121.462C11
CURRENT APPLICATION NUMBER: US/10/860,790
CURRENT FILING DATE: 2004-06-02
NUMBER OF SEQ ID NOS: 624
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 389
LENGTH: 833
TYPE: PRT
ORGANISM: Homo sapiens
US-10-860-790-389

Query Match 92.9%; Score 3663; DB 5; Length 833;
Best Local Similarity 99.9%; Pred. No. 4e-289;
Matches 696; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

14 FTHRSSVSTSTGTPTVYIGASKTPASIFGPSASHLLILFTLNFTTNLRYEENWPG 73
78 FTHRSSVSTSTGTPTVYIGASKTPASIFGPSASHLLILFTLNFTTNLRYEENWPG 137
74 SRKFNTERVLOGLRLPLFKNTSVGPLYSGCRLTLRPEKDEATGVDAICTHRPDTGP 133
138 SRKFNTERVLOGLRLPLFKNTSVGPLYSGCRLTLRPEKDEATGVDAICTHRPDTGP 197
134 GDREOLYLELSQTLHSITELGPTTLDRLSYNGFTHRSSVPTTSGVSEEPFTLNFT 193
198 GDREOLYLELSQTLHSITELGPTTLDRLSYNGFTHRSSVPTTSGVSEEPFTLNFT 257
194 INNLRYADMGQSGSLKFNITDVMKHLSPFORSSLGARYTCRVALRSYKNGAETR 253
258 INNLRYADMGQSGSLKFNITDVMKHLSPFORSSLGARYTCRVALRSYKNGAETR 317
254 VDLCTYLOPLSGRGLPIKQVFEHLSQOHTGIRLGRPLSDKSLYNGNEGPDPPT 313
318 VDLCTYLOPLSGRGLPIKQVFEHLSQOHTGIRLGRPLSDKSLYNGNEGPDPPT 377
314 TPXPATTFPLPSEATTAMGYHLKTLTNFTISNLQYSPDMGKSATFNSTEGVLOHLR 373
378 TPXPATTFPLPSEATTAMGYHLKTLTNFTISNLQYSPDMGKSATFNSTEGVLOHLR 437
374 PLFOKSMGPFYIGCOLISLRPEKGAATGVDTCTYHDPVPGGLDIQOLYWELSQTLH 433
438 PLFOKSMGPFYIGCOLISLRPEKGAATGVDTCTYHDPVPGGLDIQOLYWELSQTLH 497
434 GVTOLGFYVLDRLSLFINGAPONLSIRGEYQINFIHVMNLSNPDPSTSEYITLLRDIO 493
498 GVTOLGFYVLDRLSLFINGAPONLSIRGEYQINFIHVMNLSNPDPSTSEYITLLRDIO 557
494 DKYTTLYKSQLDHTFRFCVLTNLTMDSVLVTKALFSSNLDPDLVEOVFLDXTLNASFH 553
558 DKYTTLYKSQLDHTFRFCVLTNLTMDSVLVTKALFSSNLDPDLVEOVFLDXTLNASFH 617
554 WLGSTYQVLDIVHTEMESSYVOPTSSSTQHFYLNFTTNLPYSQDKAOPGTTNYQRNKR 613
618 WLGSTYQVLDIVHTEMESSYVOPTSSSTQHFYLNFTTNLPYSQDKAOPGTTNYQRNKR 677
614 NIEDALNOLFRRNSISYSDCOVSTFRSVPNRHHGTGVDSLCNFSLARRVDRVAIYEFP 673
678 NIEDALNOLFRRNSISYSDCOVSTFRSVPNRHHGTGVDSLCNFSLARRVDRVAIYEFP 737
674 LRMTNRGTOLQNFLLDRSSVLDVGYSPNREPLTGN 710
738 LRMTNRGTOLQNFLLDRSSVLDVGYSPNREPLTGN 774

RESULT 9
US-09-778-320-206
Sequence 206, Application US/09778320

```

; Patent No. US20010034052A1
; GENERAL INFORMATION:
; APPLICANT: Dillon, Davin C.
; APPLICANT: Day, Craig H.
; APPLICANT: Jiang, Yugu
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Mitcham, Jennifer
; APPLICANT: Wang, Tongtong
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.491C5
; CURRENT APPLICATION NUMBER: US/09/778,320
; CURRENT FILING DATE: 2001-02-06
; NUMBER OF SEQ ID NOS: 301
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO: 206
; LENGTH: 914
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-778-320-206

```

```

Query Match      92.9%; Score 3663; DB 3; Length 914;
Best Local Similarity 99.9%; Pred. No. 4,6e-289;
Matches 696; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

14  FTHRSSVSTSTPGPTVYLGASKTPASIFGSAASHLLILFTLNFTINLRYENMMPG 73
159  FTHRSSVSTSTPGPTVYLGASKTPASIFGSAASHLLILFTLNFTINLRYENMMPG 218
74  SRKFNTERVLOGLRLPFRKNTSVGPLYSGCRLTLRPEKDEAGVDAICTHRPDPGP 133
219  SRKFNTERVLOGLRLPFRKNTSVGPLYSGCRLTLRPEKDEAGVDAICTHRPDPGP 278
134  GIDREQLYLELSQTLHSITELAPYTLDRDSLVLNGFTHRSSVPTTSTGVASEEPTLNT 193
279  GIDREQLYLELSQTLHSITELAPYTLDRDSLVLNGFTHRSSVPTTSTGVASEEPTLNT 338
194  INNLRYAMDMGQPGSLKFNITDVMKHLSPFORSLGARYTGCRVIALRSYKNGAEFR 253
339  INNLRYAMDMGQPGSLKFNITDVMKHLSPFORSLGARYTGCRVIALRSYKNGAEFR 398
254  VDLCTYLOPLSGPGLPIKQVFHLSQTHGITRLGPIYSLDKDSLVLNGYNEBGPDEPPT 313
399  VDLCTYLOPLSGPGLPIKQVFHLSQTHGITRLGPIYSLDKDSLVLNGYNEBGPDEPPT 458
314  TPKPATTFPLPPLSEATTAMGYHLKTLTLNFTISNLQYSPDMKGSAATFNSTEGVLOHLR 373
459  TPKPATTFPLPPLSEATTAMGYHLKTLTLNFTISNLQYSPDMKGSAATFNSTEGVLOHLR 518
374  PLFOKSMGPFYLGCOLISLRPEKGAAGVDTTCYHDPVPGGLDIQQLYWELSQTLH 433
519  PLFOKSMGPFYLGCOLISLRPEKGAAGVDTTCYHDPVPGGLDIQQLYWELSQTLH 578
434  GTVQLGFYVLDRLSLFINGYAPQNLIRGEYOINFIHVMNLSNPPTSSEYITLLARDIQ 493
579  GTVQLGFYVLDRLSLFINGYAPQNLIRGEYOINFIHVMNLSNPPTSSEYITLLARDIQ 638
494  DKYTLTKGSQLHDTFRFCVLTNLTMDSVLVYKALFSSNLDPSELVEQVFLDKTLNLSFH 553
639  DKYTLTKGSQLHDTFRFCVLTNLTMDSVLVYKALFSSNLDPSELVEQVFLDKTLNLSFH 698
554  WLGSTYQVLDIHTHEMESVYOPTSSSSTGHFYLNFTITNLPSQDKAQPCTINYOQRNK 613
699  WLGSTYQVLDIHTHEMESVYOPTSSSSTGHFYLNFTITNLPSQDKAQPCTINYOQRNK 758
614  NIEDALNQLFRNSSIKSYFSDCOVSTFRSVPNRHHTGVDSL CNFSPILARRVDAIYEEF 673
759  NIEDALNQLFRNSSIKSYFSDCOVSTFRSVPNRHHTGVDSL CNFSPILARRVDAIYEEF 818
674  LRMTNRNGTOCONFTLDRSSVLVDGYSPNRNPLTNGS 710
819  LRMTNRNGTOCONFTLDRSSVLVDGYSPNRNPLTNGS 855

```

```

RESULT 10
US-09-910-689-206
; Sequence 206, Application US/09910689
; Patent No. US20020081609A1
; GENERAL INFORMATION:
; APPLICANT: Dillon, Davin C.
; APPLICANT: Day, Craig H.
; APPLICANT: Jiang, Yugu
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Mitcham, Jennifer
; APPLICANT: Wang, Tongtong
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.491C6
; CURRENT APPLICATION NUMBER: US/09/910,689
; CURRENT FILING DATE: 2001-07-20
; NUMBER OF SEQ ID NOS: 307
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 206
; LENGTH: 914
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-910-689-206

```

```

Query Match      92.9%; Score 3663; DB 3; Length 914;
Best Local Similarity 99.9%; Pred. No. 4,6e-289;
Matches 696; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

14  FTHRSSVSTSTPGPTVYLGASKTPASIFGSAASHLLILFTLNFTINLRYENMMPG 73
159  FTHRSSVSTSTPGPTVYLGASKTPASIFGSAASHLLILFTLNFTINLRYENMMPG 218
74  SRKFNTERVLOGLRLPFRKNTSVGPLYSGCRLTLRPEKDEAGVDAICTHRPDPGP 133
219  SRKFNTERVLOGLRLPFRKNTSVGPLYSGCRLTLRPEKDEAGVDAICTHRPDPGP 278
134  GIDREQLYLELSQTLHSITELAPYTLDRDSLVLNGFTHRSSVPTTSTGVASEEPTLNT 193
279  GIDREQLYLELSQTLHSITELAPYTLDRDSLVLNGFTHRSSVPTTSTGVASEEPTLNT 338
194  INNLRYAMDMGQPGSLKFNITDVMKHLSPFORSLGARYTGCRVIALRSYKNGAEFR 253
339  INNLRYAMDMGQPGSLKFNITDVMKHLSPFORSLGARYTGCRVIALRSYKNGAEFR 398
254  VDLCTYLOPLSGPGLPIKQVFHLSQTHGITRLGPIYSLDKDSLVLNGYNEBGPDEPPT 313
399  VDLCTYLOPLSGPGLPIKQVFHLSQTHGITRLGPIYSLDKDSLVLNGYNEBGPDEPPT 458
314  TPKPATTFPLPPLSEATTAMGYHLKTLTLNFTISNLQYSPDMKGSAATFNSTEGVLOHLR 373
459  TPKPATTFPLPPLSEATTAMGYHLKTLTLNFTISNLQYSPDMKGSAATFNSTEGVLOHLR 518
374  PLFOKSMGPFYLGCOLISLRPEKGAAGVDTTCYHDPVPGGLDIQQLYWELSQTLH 433
519  PLFOKSMGPFYLGCOLISLRPEKGAAGVDTTCYHDPVPGGLDIQQLYWELSQTLH 578
434  GTVQLGFYVLDRLSLFINGYAPQNLIRGEYOINFIHVMNLSNPPTSSEYITLLARDIQ 493
579  GTVQLGFYVLDRLSLFINGYAPQNLIRGEYOINFIHVMNLSNPPTSSEYITLLARDIQ 638
494  DKYTLTKGSQLHDTFRFCVLTNLTMDSVLVYKALFSSNLDPSELVEQVFLDKTLNLSFH 553
639  DKYTLTKGSQLHDTFRFCVLTNLTMDSVLVYKALFSSNLDPSELVEQVFLDKTLNLSFH 698
554  WLGSTYQVLDIHTHEMESVYOPTSSSSTGHFYLNFTITNLPSQDKAQPCTINYOQRNK 613
699  WLGSTYQVLDIHTHEMESVYOPTSSSSTGHFYLNFTITNLPSQDKAQPCTINYOQRNK 758
614  NIEDALNQLFRNSSIKSYFSDCOVSTFRSVPNRHHTGVDSL CNFSPILARRVDAIYEEF 673

```

Db 759 NIEDALNOLFRNSSIKSYFSDCOVSTFRSVNRRHHTGVDSLNCNPSPLARRDVAIIEEF 818
Qy 674 LRMRNGTQLONFLLDRSSVLVDGYSNNRNEPLTGN 710
Db 819 LRMRNGTQLONFLLDRSSVLVDGYFPNRNEPLTGN 855

RESULT 11

US-09-884-441-312
; Sequence 312, Application US/09884441
; Patent No. US20020119158A1
; GENERAL INFORMATION:
; APPLICANT: Algate, Paul A.
; APPLICANT: Carter, Derrick
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.462C7
; CURRENT APPLICATION NUMBER: US/09/884,441
; NUMBER OF SEQ ID NOS: 489
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 312
; LENGTH: 914
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-884-441-312

Query Match 92.9%; Score 3663; DB 3; Length 914;
Best Local Similarity 99.9%; Pred. No. 4,66-289;
Matches 696; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 14 FTHRSSVSTSTPGTPTVYLGAASKTPASIFGSAASHLLILFTLNFTITNLRYEENMPG 73
Db 159 FTHRSSVSTSTPGTPTVYLGAASKTPASIFGSAASHLLILFTLNFTITNLRYEENMPG 218
Qy 74 SRKNTTERVLOGLRLFKNTSVGLYSGCRLLLRPEKDGATGVDAICTHRPDPGP 133
Db 219 SRKNTTERVLOGLRLFKNTSVGLYSGCRLLLRPEKDGATGVDAICTHRPDPGP 278
Qy 134 GLDREQLYLBLSQTLHSITELGPTLDRDLSLVNGFTHRSSVPTTSGVSEEPFTLNFT 193
Db 279 GLDREQLYLBLSQTLHSITELGPTLDRDLSLVNGFTHRSSVPTTSGVSEEPFTLNFT 338
Qy 194 INNLRYADMGQPSGLKFNITDNVMKHLSPLFQRSSIGARYTCRVIALRSVNGAETR 253
Db 339 INNLRYADMGQPSGLKFNITDNVMKHLSPLFQRSSIGARYTCRVIALRSVNGAETR 398
Qy 254 VDLICTYLOPLSGGLPIKQVFHLSQQTGTRIGRYSLDKDSLVLNGYNEPDPDEPT 313
Db 399 VDLICTYLOPLSGGLPIKQVFHLSQQTGTRIGRYSLDKDSLVLNGYNEPDPDEPT 458
Qy 314 TPKPATTFPLPLSEATTAMGYHLKTLTNFTISNLQYSPMGKGSATFNSTEGVLOHLR 373
Db 459 TPKPATTFPLPLSEATTAMGYHLKTLTNFTISNLQYSPMGKGSATFNSTEGVLOHLR 518
Qy 374 PLFQKSMGPFYLGCOLISLRPEKGAATGVDTTCTHPRDVGGLDIQQLYMBLSQTLH 433
Db 519 PLFQKSMGPFYLGCOLISLRPEKGAATGVDTTCTHPRDVGGLDIQQLYMBLSQTLH 578
Qy 434 GVTQOLGFYVLDRLSLFNGYAPONLSIRGEYQINFHIVNNLSNPDPSTSEYITLLARDIQ 493
Db 579 GVTQOLGFYVLDRLSLFNGYAPONLSIRGEYQINFHIVNNLSNPDPSTSEYITLLARDIQ 638
Qy 494 DKVTTLTKGSQLOHDTFRFCVLTNLTMDSVLYTVKALFSSNLDPSLVEQVFLDKTLNLSFH 553
Db 639 DKVTTLTKGSQLOHDTFRFCVLTNLTMDSVLYTVKALFSSNLDPSLVEQVFLDKTLNLSFH 698
Qy 554 WLGSYQOLVDIHVTEMESVYQPTSSSTQHFYLNFTITNLRYEENMPG 613
Db 699 WLGSYQOLVDIHVTEMESVYQPTSSSTQHFYLNFTITNLRYEENMPG 758
Qy 614 NIEDALNOLFRNSSIKSYFSDCOVSTFRSVNRRHHTGVDSLNCNPSPLARRDVAIIEEF 673

Db 759 NIEDALNOLFRNSSIKSYFSDCOVSTFRSVNRRHHTGVDSLNCNPSPLARRDVAIIEEF 818
Qy 674 LRMRNGTQLONFLLDRSSVLVDGYSNNRNEPLTGN 710
Db 819 LRMRNGTQLONFLLDRSSVLVDGYFPNRNEPLTGN 855

RESULT 12

US-09-884-441-478
; Sequence 478, Application US/09884441
; Patent No. US20020119158A1
; GENERAL INFORMATION:
; APPLICANT: Algate, Paul A.
; APPLICANT: Carter, Derrick
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.462C7
; CURRENT APPLICATION NUMBER: US/09/884,441
; NUMBER OF SEQ ID NOS: 489
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 478
; LENGTH: 914
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-884-441-478

Query Match 92.9%; Score 3663; DB 3; Length 914;
Best Local Similarity 99.9%; Pred. No. 4,66-289;
Matches 696; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 14 FTHRSSVSTSTPGTPTVYLGAASKTPASIFGSAASHLLILFTLNFTITNLRYEENMPG 73
Db 159 FTHRSSVSTSTPGTPTVYLGAASKTPASIFGSAASHLLILFTLNFTITNLRYEENMPG 218
Qy 74 SRKNTTERVLOGLRLFKNTSVGLYSGCRLLLRPEKDGATGVDAICTHRPDPGP 133
Db 219 SRKNTTERVLOGLRLFKNTSVGLYSGCRLLLRPEKDGATGVDAICTHRPDPGP 278
Qy 134 GLDREQLYLBLSQTLHSITELGPTLDRDLSLVNGFTHRSSVPTTSGVSEEPFTLNFT 193
Db 279 GLDREQLYLBLSQTLHSITELGPTLDRDLSLVNGFTHRSSVPTTSGVSEEPFTLNFT 338
Qy 194 INNLRYADMGQPSGLKFNITDNVMKHLSPLFQRSSIGARYTCRVIALRSVNGAETR 253
Db 339 INNLRYADMGQPSGLKFNITDNVMKHLSPLFQRSSIGARYTCRVIALRSVNGAETR 398
Qy 254 VDLICTYLOPLSGGLPIKQVFHLSQQTGTRIGRYSLDKDSLVLNGYNEPDPDEPT 313
Db 399 VDLICTYLOPLSGGLPIKQVFHLSQQTGTRIGRYSLDKDSLVLNGYNEPDPDEPT 458
Qy 314 TPKPATTFPLPLSEATTAMGYHLKTLTNFTISNLQYSPMGKGSATFNSTEGVLOHLR 373
Db 459 TPKPATTFPLPLSEATTAMGYHLKTLTNFTISNLQYSPMGKGSATFNSTEGVLOHLR 518
Qy 374 PLFQKSMGPFYLGCOLISLRPEKGAATGVDTTCTHPRDVGGLDIQQLYMBLSQTLH 433
Db 519 PLFQKSMGPFYLGCOLISLRPEKGAATGVDTTCTHPRDVGGLDIQQLYMBLSQTLH 578
Qy 434 GVTQOLGFYVLDRLSLFNGYAPONLSIRGEYQINFHIVNNLSNPDPSTSEYITLLARDIQ 493
Db 579 GVTQOLGFYVLDRLSLFNGYAPONLSIRGEYQINFHIVNNLSNPDPSTSEYITLLARDIQ 638
Qy 494 DKVTTLTKGSQLOHDTFRFCVLTNLTMDSVLYTVKALFSSNLDPSLVEQVFLDKTLNLSFH 553
Db 639 DKVTTLTKGSQLOHDTFRFCVLTNLTMDSVLYTVKALFSSNLDPSLVEQVFLDKTLNLSFH 698
Qy 554 WLGSYQOLVDIHVTEMESVYQPTSSSTQHFYLNFTITNLRYEENMPG 613
Db 699 WLGSYQOLVDIHVTEMESVYQPTSSSTQHFYLNFTITNLRYEENMPG 758
Qy 614 NIEDALNOLFRNSSIKSYFSDCOVSTFRSVNRRHHTGVDSLNCNPSPLARRDVAIIEEF 673

```

Db      759 NIEDALNOLFRNSSISYSDCOVSTFRSVPNRHHTGVDSLGNFSLARVDRVAIYEEF 818
Qy      674 LMRTRNGTOLQNFITLDRSSVLVDGYSPNNEPLTGN 710
Db      819 LMRTRNGTOLQNFITLDRSSVLVDGYFPNNEPLTGN 855

RESULT 13
US-09-907-969-312
/ Sequence 312, Application US/09907969
/ Publication No. US20030091580A1
/ GENERAL INFORMATION:
/ APPLICANT: Mitcham, Jennifer L.
/ APPLICANT: King, Gordon E.
/ APPLICANT: Algate, Paul A.
/ APPLICANT: Flind, Steven P.
/ APPLICANT: Retter, Marc W.
/ APPLICANT: Fanger, Gary Richard
/ APPLICANT: Reed, Steven G.
/ APPLICANT: Vedvick, Thomas S.
/ APPLICANT: Carter, Darrick
/ APPLICANT: Hill, Paul
/ APPLICANT: Albone, Earl
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
/ TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
/ FILE REFERENCE: 210121.462C8
/ CURRENT APPLICATION NUMBER: US/09/907, 969
/ CURRENT FILING DATE: 2001-07-17
/ NUMBER OF SEQ ID NOS: 596
/ SOFTWARE: FaestSeq for Windows Version 4.0
/ SEQ ID NO 312
/ LENGTH: 914
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-907-969-312

Query Match      92.9%; Score 3663; DB 3; Length 914;
Best Local Similarity 99.9%; Pred. No. 4.6e-289;
Matches 696; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      14 FTHRSSVSTSTGTPTVYLGASCTPASIFGPAASHLLILFTLNFITNLRYEENMPG 73
Db      159 FTHRSSVSTSTGTPTVYLGASCTPASIFGPAASHLLILFTLNFITNLRYEENMPG 218

Qy      74 SRKFNTERVYLGILRLPLFNKTSVGPLYSGCRLLTLRPEKDGATGVDAICTHRPDTGP 133
Db      219 SRKFNTERVYLGILRLPLFNKTSVGPLYSGCRLLTLRPEKDGATGVDAICTHRPDTGP 278

Qy      134 GLDRBQLYELISQUTHSITELGPTLDRDSLRYNGCFTHRSSVPTTSIGVSEEPFTLNFT 193
Db      279 GLDRBQLYELISQUTHSITELGPTLDRDSLRYNGCFTHRSSVPTTSIGVSEEPFTLNFT 338

Qy      194 INNRVYADMGQPSLSKFNITDNVMKHLSPFORSSLGARYTGCRIYALRSVNGAETR 253
Db      339 INNRVYADMGQPSLSKFNITDNVMKHLSPFORSSLGARYTGCRIYALRSVNGAETR 398

Qy      254 VDLCTYLQPLSGGLPIKQVFHLSQQTHTGIRLGPISLDKDSLRYNGNEPDPDEPT 313
Db      399 VDLCTYLQPLSGGLPIKQVFHLSQQTHTGIRLGPISLDKDSLRYNGNEPDPDEPT 458

Qy      314 TPKRATTFPLPLSATTAMGYHLKTLTLNFTISNLQYSPDMGKSATFNSTEGVLOHLRL 373
Db      459 TPKRATTFPLPLSATTAMGYHLKTLTLNFTISNLQYSPDMGKSATFNSTEGVLOHLRL 518

Qy      374 PLFPKSSNGPFYLGCOLISLRPEKDGATGVDTCTYHPDPVPGGLDIQOLYELISQUTH 433
Db      519 PLFPKSSNGPFYLGCOLISLRPEKDGATGVDTCTYHPDPVPGGLDIQOLYELISQUTH 578

Qy      434 GVTOLGPFYVLDRLSLFINGVAPNULSIRGEYQINFHIVNMLSNPDPSTSEYITLRLDIQ 493
Db      579 GVTOLGPFYVLDRLSLFINGVAPNULSIRGEYQINFHIVNMLSNPDPSTSEYITLRLDIQ 638

```

```

Qy      494 DKVTLTKYKSQLDHTFERFCLVTNLTMDSVLYTVKALFSSNLDPSLVEQVFLDKTINASH 553
Db      639 DKVTLTKYKSQLDHTFERFCLVTNLTMDSVLYTVKALFSSNLDPSLVEQVFLDKTINASH 698

Qy      554 WLGSTYQVLDIVHTEMESVYQPTSSSTQHLYLNFITNLRYSDDKAPGTTNQRNR 613
Db      699 WLGSTYQVLDIVHTEMESVYQPTSSSTQHLYLNFITNLRYSDDKAPGTTNQRNR 758

Qy      614 NIEDALNOLFRNSSISKYSPDCOVSTFRSVPNRHHTGVDSLGNFSLARVDRVAIYEEF 673
Db      759 NIEDALNOLFRNSSISKYSPDCOVSTFRSVPNRHHTGVDSLGNFSLARVDRVAIYEEF 818

Qy      674 LMRTRNGTOLQNFITLDRSSVLVDGYSPNNEPLTGN 710
Db      819 LMRTRNGTOLQNFITLDRSSVLVDGYFPNNEPLTGN 855

RESULT 14
US-09-907-969-478
/ Sequence 478, Application US/09907969
/ Publication No. US20030091580A1
/ GENERAL INFORMATION:
/ APPLICANT: Mitcham, Jennifer L.
/ APPLICANT: King, Gordon E.
/ APPLICANT: Algate, Paul A.
/ APPLICANT: Flind, Steven P.
/ APPLICANT: Retter, Marc W.
/ APPLICANT: Fanger, Gary Richard
/ APPLICANT: Reed, Steven G.
/ APPLICANT: Vedvick, Thomas S.
/ APPLICANT: Carter, Darrick
/ APPLICANT: Hill, Paul
/ APPLICANT: Albone, Earl
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
/ TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
/ FILE REFERENCE: 210121.462C8
/ CURRENT APPLICATION NUMBER: US/09/907, 969
/ CURRENT FILING DATE: 2001-07-17
/ NUMBER OF SEQ ID NOS: 596
/ SOFTWARE: FaestSeq for Windows Version 4.0
/ SEQ ID NO 478
/ LENGTH: 914
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-907-969-478

Query Match      92.9%; Score 3663; DB 3; Length 914;
Best Local Similarity 99.9%; Pred. No. 4.6e-289;
Matches 696; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      14 FTHRSSVSTSTGTPTVYLGASCTPASIFGPAASHLLILFTLNFITNLRYEENMPG 73
Db      159 FTHRSSVSTSTGTPTVYLGASCTPASIFGPAASHLLILFTLNFITNLRYEENMPG 218

Qy      74 SRKFNTERVYLGILRLPLFNKTSVGPLYSGCRLLTLRPEKDGATGVDAICTHRPDTGP 133
Db      219 SRKFNTERVYLGILRLPLFNKTSVGPLYSGCRLLTLRPEKDGATGVDAICTHRPDTGP 278

Qy      134 GLDRBQLYELISQUTHSITELGPTLDRDSLRYNGCFTHRSSVPTTSIGVSEEPFTLNFT 193
Db      279 GLDRBQLYELISQUTHSITELGPTLDRDSLRYNGCFTHRSSVPTTSIGVSEEPFTLNFT 338

Qy      194 INNRVYADMGQPSLSKFNITDNVMKHLSPFORSSLGARYTGCRIYALRSVNGAETR 253
Db      339 INNRVYADMGQPSLSKFNITDNVMKHLSPFORSSLGARYTGCRIYALRSVNGAETR 398

Qy      254 VDLCTYLQPLSGGLPIKQVFHLSQQTHTGIRLGPISLDKDSLRYNGNEPDPDEPT 313
Db      399 VDLCTYLQPLSGGLPIKQVFHLSQQTHTGIRLGPISLDKDSLRYNGNEPDPDEPT 458

Qy      314 TPKRATTFPLPLSATTAMGYHLKTLTLNFTISNLQYSPDMGKSATFNSTEGVLOHLRL 373
Db      459 TPKRATTFPLPLSATTAMGYHLKTLTLNFTISNLQYSPDMGKSATFNSTEGVLOHLRL 518

```

```
QY 374 PLFOKSSMGPFYLGCCOLISLRPEKDGATGVDTTCTYHPDPVPGGLDIOQLYWELSQUTH 433
| | | | |
DB 519 PLFOKSSMGPFYLGCCOLISLRPEKDGATGVDTTCTYHPDPVPGGLDIOQLYWELSQUTH 578
| | | | |
QY 434 GVTQLGFYVLDRLSLFINGYAPQNLISRGXYQINFHIVNNLSNPDPTSSSEYITLLRDIO 493
| | | | |
DB 579 GVTQLGFYVLDRLSLFINGYAPQNLISRGXYQINFHIVNNLSNPDPTSSSEYITLLRDIO 638
| | | | |
QY 494 DKYTTLYKGSQLDHDFRCCLVTNLTMDSVLYTVKALFSSNLDPSLVEQVFLDKTLNLSFH 553
| | | | |
DB 639 DKYTTLYKGSQLDHDFRCCLVTNLTMDSVLYTVKALFSSNLDPSLVEQVFLDKTLNLSFH 698
| | | | |
QY 554 WLGSYQVLDIHYTEMSSYVQPTSSSTQHFYLNFTITNLPSYQDKAOPGTNYORNR 613
| | | | |
DB 699 WLGSYQVLDIHYTEMSSYVQPTSSSTQHFYLNFTITNLPSYQDKAOPGTNYORNR 758
| | | | |
QY 614 NIEDALNQLFRNSSIKSYFSDCOVSTFRSVPNRHHTGVDSLGNFSPILARRVDRVAIYEEF 673
| | | | |
DB 759 NIEDALNQLFRNSSIKSYFSDCOVSTFRSVPNRHHTGVDSLGNFSPILARRVDRVAIYEEF 818
| | | | |
QY 674 LRMTNRGTQLQNFLLDRSSVLYVDGYSBNRNEPLTGN 710
| | | | |
DB 819 LRMTNRGTQLQNFLLDRSSVLYVDGYSBNRNEPLTGN 855
| | | | |
```

RESULT 15

```
US-09-827-271-312
; Sequence 312, Application US/09827271
; Publication No. US20030165504A1
; GENERAL INFORMATION:
; APPLICANT: Retter, Marc W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.462C6
; CURRENT APPLICATION NUMBER: US/09/827,271
; CURRENT FILING DATE: 2001-04-04
; NUMBER OF SEQ ID NOS: 461
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 312
; LENGTH: 914
; TYPE: PR1
; ORGANISM: Homo sapien
US-09-827-271-312
```

Query Match 92.9%; Score 3663; DB 3; Length 914;

Best Local Similarity 99.9%; Pred. No. 4.6e-289;

Matches 696; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 14 FTHRSSVSTSTPGTPTVYLGASKTSPASIFGPSASHLLILFTLNFTITNLRYEENMPG 73
| | | | |
DB 159 FTHRSSVSTSTPGTPTVYLGASKTSPASIFGPSASHLLILFTLNFTITNLRYEENMPG 218
| | | | |
QY 74 SRKNTTERVYQGLLRPLFKNTSVGPLYSGCRLTLRLPEKDGATGVDAICTHRRDPTGP 133
| | | | |
DB 219 SRKNTTERVYQGLLRPLFKNTSVGPLYSGCRLTLRLPEKDGATGVDAICTHRRDPTGP 278
| | | | |
QY 134 GLDREQLYELISQUTHSITEGPTLDRDSL YVNGFTHRSSVPTTSGVSEBPTLNFT 193
| | | | |
DB 279 GLDREQLYELISQUTHSITEGPTLDRDSL YVNGFTHRSSVPTTSGVSEBPTLNFT 338
| | | | |
QY 194 INNIRYADMGQPSLKFENITDNVMKHLSPLFQSSIGARYTGCRVIALRSVNGAETR 253
| | | | |
DB 339 INNIRYADMGQPSLKFENITDNVMKHLSPLFQSSIGARYTGCRVIALRSVNGAETR 398
| | | | |
QY 254 VDLICTYLOPLSGGLPIKOVFHELISOOTGITRLGYSLDKSL YLNGVNEBGPDEPPT 313
| | | | |
DB 399 VDLICTYLOPLSGGLPIKOVFHELISOOTGITRLGYSLDKSL YLNGVNEBGPDEPPT 458
| | | | |
QY 314 TPXPATFTPLPLSEATTAMGYHLKTLNLNFTISNLQYSPDMGKSATFNSSTEGVLQHLR 373
| | | | |
DB 459 TPXPATFTPLPLSEATTAMGYHLKTLNLNFTISNLQYSPDMGKSATFNSSTEGVLQHLR 518
| | | | |
```

```
QY 374 PLFOKSSMGPFYLGCCOLISLRPEKDGATGVDTTCTYHPDPVPGGLDIOQLYWELSQUTH 433
| | | | |
DB 519 PLFOKSSMGPFYLGCCOLISLRPEKDGATGVDTTCTYHPDPVPGGLDIOQLYWELSQUTH 578
| | | | |
QY 434 GVTQLGFYVLDRLSLFINGYAPQNLISRGXYQINFHIVNNLSNPDPTSSSEYITLLRDIO 493
| | | | |
DB 579 GVTQLGFYVLDRLSLFINGYAPQNLISRGXYQINFHIVNNLSNPDPTSSSEYITLLRDIO 638
| | | | |
QY 494 DKYTTLYKGSQLDHDFRCCLVTNLTMDSVLYTVKALFSSNLDPSLVEQVFLDKTLNLSFH 553
| | | | |
DB 639 DKYTTLYKGSQLDHDFRCCLVTNLTMDSVLYTVKALFSSNLDPSLVEQVFLDKTLNLSFH 698
| | | | |
QY 554 WLGSYQVLDIHYTEMSSYVQPTSSSTQHFYLNFTITNLPSYQDKAOPGTNYORNR 613
| | | | |
DB 699 WLGSYQVLDIHYTEMSSYVQPTSSSTQHFYLNFTITNLPSYQDKAOPGTNYORNR 758
| | | | |
QY 614 NIEDALNQLFRNSSIKSYFSDCOVSTFRSVPNRHHTGVDSLGNFSPILARRVDRVAIYEEF 673
| | | | |
DB 759 NIEDALNQLFRNSSIKSYFSDCOVSTFRSVPNRHHTGVDSLGNFSPILARRVDRVAIYEEF 818
| | | | |
QY 674 LRMTNRGTQLQNFLLDRSSVLYVDGYSBNRNEPLTGN 710
| | | | |
DB 819 LRMTNRGTQLQNFLLDRSSVLYVDGYSBNRNEPLTGN 855
| | | | |
```

Search completed: October 14, 2006, 04:09:19

Job time : 189 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: October 14, 2006, 04:06:25 : Search time 40 Seconds

(without alignments)
1484.879 Million cell updates/sec

Title: US-10-687-035-1

Perfect score: 3945

Sequence: 1 AAGPARARRRTKLFTHRSSV.....QKLISEEDLMMHTGHHHHH 748

Scoring table:

BLOSUM62

Searched: 295242 seqs, 79405279 residues

Total number of hits satisfying chosen parameters: 295242

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA New:
1: /EMC_Ceiera_SIDS3/prodata/1/pubppaa/US09_NEW_PUB pep.*
2: /EMC_Ceiera_SIDS3/prodata/1/pubppaa/US06_NEW_PUB pep.*
3: /EMC_Ceiera_SIDS3/prodata/1/pubppaa/US07_NEW_PUB pep.*
4: /EMC_Ceiera_SIDS3/prodata/1/pubppaa/US08_NEW_PUB pep.*
5: /EMC_Ceiera_SIDS3/prodata/1/pubppaa/PCT_NEW_PUB pep.*
6: /EMC_Ceiera_SIDS3/prodata/1/pubppaa/US10_NEW_PUB pep.*
7: /EMC_Ceiera_SIDS3/prodata/1/pubppaa/US11_NEW_PUB pep.*
8: /EMC_Ceiera_SIDS3/prodata/1/pubppaa/US60_NEW_PUB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----------------------|--------------------|
| 1 | 3649 | 92.5 | 22152 | US-10-544-944-1 | Sequence 1, Appli |
| 2 | 3642 | 92.3 | 1148 | US-11-105-233-195 | Sequence 135, App |
| 3 | 3642 | 92.3 | 1148 | US-11-226-554-129 | Sequence 129, App |
| 4 | 3642 | 92.3 | 1148 | US-11-248-718-129 | Sequence 129, App |
| 5 | 182 | 4.6 | 33 | US-11-134-871-1509 | Sequence 1509, Ap |
| 6 | 161.5 | 4.1 | 396 | US-10-541-657-2 | Sequence 2, Appli |
| 7 | 161.5 | 4.1 | 396 | US-11-327-896A-2 | Sequence 2, Appli |
| 8 | 158.5 | 4.0 | 288 | US-10-345-557-48 | Sequence 48, Appli |
| 9 | 146 | 3.7 | 625 | US-10-545-557-40 | Sequence 40, Appli |
| 10 | 136 | 3.4 | 542 | US-11-056-355B-82537 | Sequence 82537, A |
| 11 | 136 | 3.4 | 605 | US-11-056-355B-82536 | Sequence 82536, A |
| 12 | 136 | 3.4 | 609 | US-11-056-355B-82535 | Sequence 82535, A |
| 13 | 132 | 3.3 | 774 | US-11-123-682-34 | Sequence 34, Appli |
| 14 | 130 | 3.3 | 401 | US-11-234-587-355 | Sequence 355, App |
| 15 | 130 | 3.3 | 401 | US-11-234-482-15 | Sequence 15, Appli |
| 16 | 121.5 | 3.0 | 1475 | US-10-527-100-33 | Sequence 33, Appli |
| 17 | 118.5 | 3.0 | 270 | US-11-334-081-36 | Sequence 36, Appli |
| 18 | 117 | 3.0 | 21 | US-11-332-088-2 | Sequence 2, Appli |
| 19 | 117 | 3.0 | 21 | US-11-348-129-30 | Sequence 30, Appli |
| 20 | 114.5 | 2.9 | 211 | US-11-334-081-54 | Sequence 44, Appli |
| 21 | 114.5 | 2.9 | 233 | US-11-334-081-55 | Sequence 55, Appli |
| 22 | 114.5 | 2.9 | 273 | US-11-334-081-40 | Sequence 40, Appli |
| 23 | 114 | 2.9 | 489 | US-10-498-697-1 | Sequence 1, Appli |
| 24 | 113.5 | 2.9 | 539 | US-11-123-692-37 | Sequence 37, Appli |
| 25 | 113.5 | 2.9 | 539 | US-11-123-692-42 | Sequence 42, Appli |

ALIGNMENTS

```
RESULT 1
US-10-544-944-1
Sequence 1, Appli Application US/10544944
Publication No. US20060134120A1
GENERAL INFORMATION:
APPLICANT: Diamandis, Eleftherios P.
TITLE OR INVENTION: Multiple Marker Assay for Detection of Ovarian Cancer
FILE REFERENCE: 11757.008805W0
CURRENT APPLICATION NUMBER: US/10/544,944
PRIOR FILING DATE: 2005-08-09
PRIOR APPLICATION NUMBER: PCT/CA2004/000281
PRIOR FILING DATE: 2004-02-26
PRIOR APPLICATION NUMBER: US 60/450,406
NUMBER OF SEQ ID NOS: 22
SOFTWARE: PatentIn version 3.2
SEQ ID NO 1
LENGTH: 22152
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (13877)..(13878)
OTHER INFORMATION: Xaa can be any naturally occurring amino acid
FEATURE:
NAME/KEY: misc feature
LOCATION: (13880)..(13880)
OTHER INFORMATION: Xaa can be any naturally occurring amino acid
FEATURE:
NAME/KEY: misc feature
LOCATION: (13890)..(13891)
OTHER INFORMATION: Xaa can be any naturally occurring amino acid
FEATURE:
NAME/KEY: misc feature
LOCATION: (13893)..(13893)
OTHER INFORMATION: Xaa can be any naturally occurring amino acid
FEATURE:
NAME/KEY: misc feature
LOCATION: (13903)..(13903)
OTHER INFORMATION: Xaa can be any naturally occurring amino acid
FEATURE:
NAME/KEY: misc feature
LOCATION: (13914)..(13914)
OTHER INFORMATION: Xaa can be any naturally occurring amino acid
```



```
; LOCATION: (16046)..(16046)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (16051)..(16051)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
; FEATURE:
; NAME/KEY: misc_feature
```

```
Query Match      92.5%; Score 3649; DB 6; Length 22152;
Best Local Similarity 99.7%; Pred. No. 1e-263;
Matches 695; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 14 FTHRSSVSTSTGTPTVYLGASKTPTASIFGPAASHLLILFTLNFTITNLRYEENMPG 73
DB 21397 FTHRSSVSTSTGTPTVYLGASKTPTASIFGPAASHLLILFTLNFTITNLRYEENMPG 21456
QY 74 SRKFTTERVLOGLRLPFRNTSVGPLYSGCRLLTLRPEKGEATGVDAICTHRPDPGTP 133
DB 21457 SRKFTTERVLOGLRLPFRNTSVGPLYSGCRLLTLRPEKGEATGVDAICTHRPDPGTP 21516
QY 134 GLDREQLYLELSQTLHSITELGPTTLDRLSLYNGFTHRSSVPTTSTGVSEEPFTLNFT 193
DB 21517 GLDREQLYLELSQTLHSITELGPTTLDRLSLYNGFTHRSSVPTTSTGVSEEPFTLNFT 21576
QY 194 INNLRYMADMGQSGSLKFNITDVMKHLSPLFORSISLGARYGCRVIALRSYKNGAETR 253
DB 21577 INNLRYMADMGQSGSLKFNITDVMKHLSPLFORSISLGARYGCRVIALRSYKNGAETR 21636
QY 254 VDLICTYLOPLSGRGLPIKQVPHFELSQQTHGTRLGPSLDKDSLVLNGYNEGPDPEPT 313
DB 21637 VDLICTYLOPLSGRGLPIKQVPHFELSQQTHGTRLGPSLDKDSLVLNGYNEGPDPEPT 21696
QY 314 TPKPATTFPLPULSEATTAMGYHLKTLTNFTISNLQYSPDMGKGSATFNSTEGVLOHLR 373
DB 21697 TPKPATTFPLPULSEATTAMGYHLKTLTNFTISNLQYSPDMGKGSATFNSTEGVLOHLR 21756
QY 374 PLFOKSSMGPFYLGCOLISLRPEKGAATGVDTCTYHPDPVPGDIOQLYWEELSOLTH 433
DB 21757 PLFOKSSMGPFYLGCOLISLRPEKGAATGVDTCTYHPDPVPGDIOQLYWEELSOLTH 21816
QY 434 GVTQLGFVYLDRLSLFINGYAPONLSIRGEYQINFHIVMNLNSNPPTSEYITLLRDIO 493
DB 21817 GVTQLGFVYLDRLSLFINGYAPONLSIRGEYQINFHIVMNLNSNPPTSEYITLLRDIO 21876
QY 494 DKVTTLYKGSQLDHTRFCLVTNLTMDSVLVTYKALFSSNLDPDLVEQVFLDKTLNASFH 553
DB 21877 DKVTTLYKGSQLDHTRFCLVTNLTMDSVLVTYKALFSSNLDPDLVEQVFLDKTLNASFH 21936
QY 554 WLGSTYQLVDIHYTEMESVYQPTSSSTQHFYLNFTITNLPSODKAPGTTNYORNR 613
DB 21937 WLGSTYQLVDIHYTEMESVYQPTSSSTQHFYLNFTITNLPSODKAPGTTNYORNR 21996
QY 614 NIEDALNOLFRNSSISYSPDCOVSTFRSVPNRHHGVDSLGNFSLARVDVAIYEER 673
DB 21997 NIEDALNOLFRNSSISYSPDCOVSTFRSVPNRHHGVDSLGNFSLARVDVAIYEER 22056
QY 674 LRMTNGTQLQNFLLDRSSVLVDGYSPPNRNEPLTGN 710
DB 22057 LRMTNGTQLQNFLLDRSSVLVDGYSPPNRNEPLTGN 22093
```

```
RESULT 2
US-11-105-233-195
; Sequence 195, Application US/11105233
; Publication No. US20060134653A1
```

```
; GENERAL INFORMATION:
```

```
; APPLICANT: Thiagalingam et al
```

```
; TITLE OF INVENTION: Differential Expression of Genes in MSI
```

```
; FILE REFERENCE: 1657/2001
```

```
; CURRENT APPLICATION NUMBER: US/11/105,233
```

```
; CURRENT FILING DATE: 2005-04-13
```

```
; NUMBER OF SEQ ID NOS: 202
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 195
; LENGTH: 1148
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-105-233-195
```

```
Query Match      92.3%; Score 3642; DB 7; Length 1148;
Best Local Similarity 99.6%; Pred. No. 4.6e-265;
Matches 694; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 14 FTHRSSVSTSTGTPTVYLGASKTPTASIFGPAASHLLILFTLNFTITNLRYEENMPG 73
DB 393 FTHRSSVSTSTGTPTVYLGASKTPTASIFGPAASHLLILFTLNFTITNLRYEENMPG 452
QY 74 SRKFTTERVLOGLRLPFRNTSVGPLYSGCRLLTLRPEKGEATGVDAICTHRPDPGTP 133
DB 453 SRKFTTERVLOGLRLPFRNTSVGPLYSGCRLLTLRPEKGEATGVDAICTHRPDPGTP 512
QY 134 GLDREQLYLELSQTLHSITELGPTTLDRLSLYNGFTHRSSVPTTSTGVSEEPFTLNFT 193
DB 513 GLDREQLYLELSQTLHSITELGPTTLDRLSLYNGFTHRSSVPTTSTGVSEEPFTLNFT 572
QY 194 INNLRYMADMGQSGSLKFNITDVMKHLSPLFORSISLGARYGCRVIALRSYKNGAETR 253
DB 573 INNLRYMADMGQSGSLKFNITDVMKHLSPLFORSISLGARYGCRVIALRSYKNGAETR 632
QY 254 VDLICTYLOPLSGRGLPIKQVPHFELSQQTHGTRLGPSLDKDSLVLNGYNEGPDPEPT 313
DB 633 VDLICTYLOPLSGRGLPIKQVPHFELSQQTHGTRLGPSLDKDSLVLNGYNEGPDPEPT 692
QY 314 TPKPATTFPLPULSEATTAMGYHLKTLTNFTISNLQYSPDMGKGSATFNSTEGVLOHLR 373
DB 693 TPKPATTFPLPULSEATTAMGYHLKTLTNFTISNLQYSPDMGKGSATFNSTEGVLOHLR 752
QY 374 PLFOKSSMGPFYLGCOLISLRPEKGAATGVDTCTYHPDPVPGDIOQLYWEELSOLTH 433
DB 753 PLFOKSSMGPFYLGCOLISLRPEKGAATGVDTCTYHPDPVPGDIOQLYWEELSOLTH 812
QY 434 GVTQLGFVYLDRLSLFINGYAPONLSIRGEYQINFHIVMNLNSNPPTSEYITLLRDIO 493
DB 813 GVTQLGFVYLDRLSLFINGYAPONLSIRGEYQINFHIVMNLNSNPPTSEYITLLRDIO 872
QY 494 DKVTTLYKGSQLDHTRFCLVTNLTMDSVLVTYKALFSSNLDPDLVEQVFLDKTLNASFH 553
DB 873 DKVTTLYKGSQLDHTRFCLVTNLTMDSVLVTYKALFSSNLDPDLVEQVFLDKTLNASFH 932
QY 554 WLGSTYQLVDIHYTEMESVYQPTSSSTQHFYLNFTITNLPSODKAPGTTNYORNR 613
DB 933 WLGSTYQLVDIHYTEMESVYQPTSSSTQHFYLNFTITNLPSODKAPGTTNYORNR 992
QY 614 NIEDALNOLFRNSSISYSPDCOVSTFRSVPNRHHGVDSLGNFSLARVDVAIYEER 673
DB 993 NIEDALNOLFRNSSISYSPDCOVSTFRSVPNRHHGVDSLGNFSLARVDVAIYEER 1052
QY 674 LRMTNGTQLQNFLLDRSSVLVDGYSPPNRNEPLTGN 710
DB 1053 LRMTNGTQLQNFLLDRSSVLVDGYSPPNRNEPLTGN 1089
```

```
RESULT 3
US-11-226-554-129
; Sequence 129, Application US/11226554
; Publication No. US20060147373A1
```

```
; GENERAL INFORMATION:
```

```
; APPLICANT: Cairns, Belinda
```

```
; APPLICANT: Chen, Ruihuan
```

```
; APPLICANT: Frantz, Gretchen
```

```
; APPLICANT: Hillan, Kenneth J.
```

```
; APPLICANT: Koeppe, Hartmut
```

```
; APPLICANT: Phillips, Heidi S.
```

```
; APPLICANT: Polakis, Paul
```

```

; APPLICANT: Spencer, Susan D.
; APPLICANT: Smith, Victoria
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wu, Thomas D.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Compositions and Methods for the Diagnosis and
; TITLE OF INVENTION: Treatment of Tumor
; FILE REFERENCE: P5001R1C1
; CURRENT APPLICATION NUMBER: US/11/226,554
; CURRENT FILING DATE: 2005-09-13
; PRIOR APPLICATION NUMBER: US 10/177,488
; PRIOR FILING DATE: 2002-06-19
; PRIOR APPLICATION NUMBER: US 60/366,869
; PRIOR FILING DATE: 2002-03-20
; NUMBER OF SEQ ID NOS: 154
; SEQ ID NO 129
; LENGTH: 1148
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-226-554-129

```

Query Match 92.3%; Score 3642; DB 7; Length 1148;

Best Local Similarity 99.6%; Pred. No. 4,6e-265;

Matches 694; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

QY 14 FTHRSVSTSTPGTPTVYLGASKTPASIFGSAASHLLILFTLNTTINLRYEENMPG 73
DB 393 FTHRSVSTSTPGTPTVYLGASKTPASIFGSAASHLLILFTLNTTINLRYEENMPG 452
QY 74 SRKFTTERVLOGLRLPLFNKTSVGLVSGGRLTLRLPKDGAATGVDALCTHRPDPGP 133
DB 453 SRKFTTERVLOGLRLPLFNKTSVGLVSGGRLTLRLPKDGAATGVDALCTHRPDPGP 512
QY 134 GLDREOLYELSQLTHSITELGPTTLDROSLVYNGFTHRSVPTTSGVVSSEPTLNT 193
DB 513 GLDREOLYELSQLTHSITELGPTTLDROSLVYNGFTHRSVPTTSGVVSSEPTLNT 572
QY 194 INNLRYADMGOQPSLKFNITDVMKHLSPLFORSISGARVYGCRIYALRSYNGAETR 253
DB 573 INNLRYADMGOQPSLKFNITDVMKHLSPLFORSISGARVYGCRIYALRSYNGAETR 632
QY 254 VDLCTYLOPLSGRGLPIKOVFHELSQOHTGIRLDPYSIDKSLVYNGNEPDPBPT 313
DB 633 VDLCTYLOPLSGRGLPIKOVFHELSQOHTGIRLDPYSIDKSLVYNGNEPDPBPT 692
QY 314 TPKPATTPPLPSEATTAMGYHLKTLTINFTISNLOYSPPMGKGSATFNSTEGVLOHLR 373
DB 693 TPKPATTPPLPSEATTAMGYHLKTLTINFTISNLOYSPPMGKGSATFNSTEGVLOHLR 752
QY 374 PLFOKSSMGPFYLGCOLISLRPEKGAATGVDTCYHPDPVGPGLDIOQLYELSQLTH 433
DB 753 PLFOKSSMGPFYLGCOLISLRPEKGAATGVDTCYHPDPVGPGLDIOQLYELSQLTH 812
QY 434 GTVQOLGPTVLDROSLFINGYAPQNLSTRGYQINFNHVMNLSNPDPSTSEYITLLRDIO 493
DB 813 GTVQOLGPTVLDROSLFINGYAPQNLSTRGYQINFNHVMNLSNPDPSTSEYITLLRDIO 872
QY 494 DKYTLTKGSQLDHTFPFCLVNTLMDSVLYVYKALFSSNLDPSLVEQVFLDKTLNLSFH 553
DB 873 DKYTLTKGSQLDHTFPFCLVNTLMDSVLYVYKALFSSNLDPSLVEQVFLDKTLNLSFH 932
QY 554 WLGSYVQVLDIHYTEMSSVYQPTSSSTGHFYINFTITNLPSQDRAQPGTINYQNRK 613
DB 933 WLGSYVQVLDIHYTEMSSVYQPTSSSTGHFYINFTITNLPSQDRAQPGTINYQNRK 992
QY 614 NIBDALNQLFRNSISKYSFSDCOVSTRSPVNRHHTGVDSICNSPLARVDRAIYEEF 673
DB 993 NIBDALNQLFRNSISKYSFSDCOVSTRSPVNRHHTGVDSICNSPLARVDRAIYEEF 1052
QY 674 LRMTRNGTOLONFTLDRSSVYLDGYSPPRNEPLTNGS 710
DB 1053 LRMTRNGTOLONFTLDRSSVYLDGYSPPRNEPLTNGS 1089

```

RESULT 4
US-11-248-718-129

Sequence 129, Application US/11248718

Publication No. US2006016097A1

GENERAL INFORMATION:

APPLICANT: Cairns, Belinda

APPLICANT: Chen, Ruihuan

APPLICANT: Frantz, Gretchen

APPLICANT: Hillan, Kenneth J.

APPLICANT: Koepfen, Hartmut

APPLICANT: Phillips, Heidi S.

APPLICANT: Polakis, Paul

APPLICANT: Spencer, Susan D.

APPLICANT: Smith, Victoria

APPLICANT: Williams, P. Mickey

APPLICANT: Wu, Thomas D.

APPLICANT: Zhang, Zemin

TITLE OF INVENTION: Compositions and Methods for the Diagnosis and

TITLE OF INVENTION: Treatment of Tumor

FILE REFERENCE: P5001R1 US

CURRENT APPLICATION NUMBER: US/11/248,718

CURRENT FILING DATE: 2005-10-11

PRIOR APPLICATION NUMBER: US/10/177,488

PRIOR FILING DATE: 2002-06-19

PRIOR APPLICATION NUMBER: US 60/299,500

PRIOR FILING DATE: 2001-06-20

PRIOR APPLICATION NUMBER: US 60/300,880

PRIOR FILING DATE: 2001-06-25

PRIOR APPLICATION NUMBER: US 60/301,880

PRIOR FILING DATE: 2001-06-29

PRIOR APPLICATION NUMBER: US 60/304,813

PRIOR FILING DATE: 2001-07-11

PRIOR APPLICATION NUMBER: US 60/312,312

PRIOR FILING DATE: 2001-08-13

PRIOR APPLICATION NUMBER: US 60/314,280

PRIOR FILING DATE: 2001-08-22

PRIOR APPLICATION NUMBER: US 60/339,227

PRIOR FILING DATE: 2001-10-19

PRIOR APPLICATION NUMBER: US 60/323,268

PRIOR FILING DATE: 2001-09-18

PRIOR APPLICATION NUMBER: US 60/336,827

PRIOR FILING DATE: 2001-11-07

Remaining Prior Application data removed - See File Wrapper or PALM.

NUMBER OF SEQ ID NOS: 154

SEQ ID NO 129

LENGTH: 1148

TYPE: PRT

ORGANISM: Homo Sapien

US-11-248-718-129

Query Match 92.3%; Score 3642; DB 7; Length 1148;

Best Local Similarity 99.6%; Pred. No. 4,6e-265;

Matches 694; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

QY 14 FTHRSVSTSTPGTPTVYLGASKTPASIFGSAASHLLILFTLNTTINLRYEENMPG 73
DB 393 FTHRSVSTSTPGTPTVYLGASKTPASIFGSAASHLLILFTLNTTINLRYEENMPG 452
QY 74 SRKFTTERVLOGLRLPLFNKTSVGLVSGGRLTLRLPKDGAATGVDALCTHRPDPGP 133
DB 453 SRKFTTERVLOGLRLPLFNKTSVGLVSGGRLTLRLPKDGAATGVDALCTHRPDPGP 512
QY 134 GLDREOLYELSQLTHSITELGPTTLDROSLVYNGFTHRSVPTTSGVVSSEPTLNT 193
DB 513 GLDREOLYELSQLTHSITELGPTTLDROSLVYNGFTHRSVPTTSGVVSSEPTLNT 572
QY 194 INNLRYADMGOQPSLKFNITDVMKHLSPLFORSISGARVYGCRIYALRSYNGAETR 253
DB 573 INNLRYADMGOQPSLKFNITDVMKHLSPLFORSISGARVYGCRIYALRSYNGAETR 632
QY 254 VDLCTYLOPLSGRGLPIKOVFHELSQOHTGIRLDPYSIDKSLVYNGNEPDPBPT 313

```



```

; APPLICANT: ANOSYS
; TITLE OF INVENTION: METHODS AND COMPOUNDS FOR RAISING ANTIBODIES AND FOR SCREENING AN
; TITLE OF INVENTION: REPERTOIRES
; FILE REFERENCE: 3665-158
; CURRENT APPLICATION NUMBER: US/10/545,557
; CURRENT FILING DATE: 2005-08-15
; PRIOR APPLICATION NUMBER: PCT/IB2004/000888
; PRIOR FILING DATE: 2004-02-13
; PRIOR APPLICATION NUMBER: US60/447,291
; PRIOR FILING DATE: 2003-02-14
; NUMBER OF SEQ ID NOS: 86
; SOFTWARE: Patent version 3.1
; SEQ ID NO: 48
; LENGTH: 288
; TYPE: PRT
; ORGANISM: artificial sequence
; FEATURE:
; OTHER INFORMATION: LS-GFP-Myc/His
US-10-545-557-48

Query Match          4.0%; Score 158.5; DB 6; Length 288;
Best Local Similarity 24.6%; Pred. No. 0.00032;
Matches 81; Conservative 30; Mismatches 111; Indels 107; Gaps 14;

QY 443 LDRSLFINGVAPONLSIRGEYO-----INPHIVNMNLSNPDPTSEYITLIRDIQ 493
DB 44 LGDGD---VNGH---KFSVSGEGSGDATYGLTLKFTCTTGKLPVMPPT----- 85
QY 494 DRYTTLTKYSQI-----HDTFRFLVNLNLTMDSLVLYVKALFSSNLDPSLVEQVF 543
DB 86 -LVTLTYGVQCFSRYPDMKQIDFPK-----SAPBEGYVERITF-----F 126
QY 544 LD-----KTLNASHFWMLGSTYQLVDIHVTEMESSVYOPTSSSTQHFYLNFTITNLPYSQD 599
DB 127 KDDGNKYT-RAEYKFEQDTL---VRIELKGIDEFEDGNILGHKLEYNVNSHNYIMAD 181
QY 600 KAPQPTTNYORNRKNTEDALNQLFRNSISIKYPSDCQVSTFSPNRHHTGVDSLGNFSP 659
DB 182 KQNGKIKANFKIRHNIEDGSVQLADHYQQNTPIGDPFV---LLPNHYLSTQSALSXP 237
QY 660 LARVRVVAIYEFLMKTNRNGTOLQNTLDRSSVLDGYSPPNNEPLTGNASADIQSGGR 719
DB 238 NEKRDHNVLL--EF--VTAGITL-----GM 259
QY 720 SLEGPREFEOKLISEEDLNMTGHHHHH 748
DB 260 DELYKRFEOKLISEEDLNMTGHHHHH 288

RESULT 9
US-10-545-557-40
; Sequence 40, Application US/10545557
; Publication No. US20060222654A1
; GENERAL INFORMATION:
; APPLICANT: ANOSYS
; TITLE OF INVENTION: METHODS AND COMPOUNDS FOR RAISING ANTIBODIES AND FOR SCREENING AN
; TITLE OF INVENTION: REPERTOIRES
; FILE REFERENCE: 3665-158
; CURRENT APPLICATION NUMBER: US/10/545,557
; CURRENT FILING DATE: 2005-08-15
; PRIOR APPLICATION NUMBER: PCT/IB2004/000888
; PRIOR FILING DATE: 2004-02-13
; PRIOR APPLICATION NUMBER: US60/447,291
; PRIOR FILING DATE: 2003-02-14
; NUMBER OF SEQ ID NOS: 86
; SOFTWARE: Patent version 3.1
; SEQ ID NO: 40
; LENGTH: 625
; TYPE: PRT
; ORGANISM: artificial sequence
; FEATURE:
; OTHER INFORMATION: LS-GFP-C1/C2
US-10-545-557-40

```

```

Query Match          3.7%; Score 146; DB 6; Length 625;
Best Local Similarity 19.2%; Pred. No. 0.0086;
Matches 124; Conservative 67; Mismatches 178; Indels 276; Gaps 31;

QY 191 NFTINLRVYADMQPSGLKFNITDYNWKILLSPLFORSSIGARYTCRVIALRSYKGA 250
DB 170 NYNSHNYIMADK-QKXGDKIKNFK---IRH-----NIEDGS 201
QY 251 ETRVDLLCTTYQ--PLSGPGLPIQVFEHLSQOT-----HGTRLRGYS 292
DB 202 ---VOLADHYQQNTPI-GDGPVLLPDNHYLSTQSALSKDPNEKDHVLLFVYTAAG-IT 256
QY 293 LDKSLYLINGYNEGQDPDPTTPKATFPLPLSATTAMGY---HLKTLT-----NF 343
DB 257 LGMDELYR-----PSTYCTLKGYAGNHCHCTKCVBLPMNG 295
QY 344 TISNLQSPDMGKSGATFNSTEGVLOHLRLPLFOKSMGPFYLGQ-----LISRPKD 398
DB 296 NIANSQIA-----ASSVRYTFGLQHWVPELARLN----- 325
QY 399 GAATGVDTTCYHPDPVPGGLDIQOL--YWEISQUTHGVTOGLFYVLDRLDSLFIGVAPQ 456
DB 326 --RAGMVAWMTFSSNDNPMIQVNLRRMWVTVGVTOGASRLASHRYK----- 372
QY 457 NLSIRGEYQINPH-----IYVNMNLSN-----PDPTSEYITLIRDIQDK 495
DB 373 --AFVVAISLNGHEPDFIHDVKKKKEFVGNNKNAHVNLFEYVEAQVYLVYPTSGHT 430
QY 496 VTTL---YKGSQIADTFRCFLVTLNLTMDSLVLYVKA-----LFSSNLDPSLVEQVF 544
DB 431 ACTLRFELLGCELNGCANPLGLKNNISIPDKQITASSYKTMGLHFSNN--PSYAR--L 485
QY 545 DXTLNASFWMLGSTY---QVLDIHV---TEMESSVYOPTSSSTQHFYLNFTITNLPYSQ 598
DB 486 DKQGNFN-AMVAGSYGNDQWLQVDLGSSKEYVGIITQGARNGSQVFAVSRYVA---YSN 541
QY 599 DKAQ-----PGTTYORNRKNTEDALNQLFRNSISIKYPSDCQVSTFSPRSV 643
DB 542 DSANWTEYODPRTGSKTIFPGWMDHSHKKN-----LFETPIRLARY-----VRL 586
QY 644 PRHHTGVDSLGNFSLARVRVAIYEFLMKTNRNGTOLQNTLDRSSVLDGYSPPN 703
DB 587 PVAMH-----NRIALRLLELGC----- 603
QY 704 EPLTGNASADIQSGGRSLEGPREFEOKLISEEDLNMTGHHHHH 748
DB 604 -----FEOKLISEEDLNMTGHHHHH 625

RESULT 10
US-11-056-355B-82537
; Sequence 82537, Application US/11056355B
; Publication No. US20060150283A1
; GENERAL INFORMATION:
; APPLICANT: Brover, Vyacheslav
; APPLICANT: Alexandrov, Nikolai
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
; TITLE OF INVENTION: Polypeptides Encoded Thereby
; FILE REFERENCE: 2750-1590PUS2
; CURRENT APPLICATION NUMBER: US/11/056,355B
; CURRENT FILING DATE: 2005-02-14
; PRIOR APPLICATION NUMBER: 60/544,190
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 119966
; SEQ ID NO 82537
; LENGTH: 542
; TYPE: prt
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide
; LOCATION: (1) - (542)
; OTHER INFORMATION: Cereas Seq. ID no. 12666070
US-11-056-355B-82537

```

US-11-056-355B-82537

Query Match 3.4%; Score 136; DB 7; Length 542;

Best Local Similarity 20.6%; Pred. No. 0.039;

Matches 99; Conservative 67; Mismatches 185; Indels 130; Gaps 20;

113 KDGATGVDAICTHRPDPGGLDREQLYELSQLTHSITELGPTTLDRDLSLVNGFTHR 172
DB RDEEVOG-DTVIERND-----QSKLETAKL-----RIYDQWIDVNLRESEYEN 134
173 SSVPTTSGVSEEPPTLNTNNLRVYADMGQPSL-KFNITDVMKHLSPFORSS- 230
DB SRIPTMKFGTAKDAPFRDLTINSLFYNNINGAVEDLTERGIDDLKSGKITVPLPAKATF 194
231 -----LGAATGCRVIALRSVNGAETRDLLCTYLQPLSGFGLPIKOVFHEL 278
DB 195 LDDPLRLVLAVERGARFGFTLDEELKEAASSEVERVAL-----GEKISRERIGNEI 245
QY 279 SQQTG-----ITRLGPYSLDKSLYLNGVNEPDPED-----PTTPKP 317
DB 246 DLMISNGPVSATVITLSDKLFSVVPALPSSAEPSPENCGLSOSYLEAMMSLKTTPR 305
QY 318 -----ATFLP-----PLSEATTAMGYHLK 337
DB 306 GKFSGEQRRLALYAAAFPLPRKTVYKDTGKSIPIVNNHIFKFSMKRKTSDAETVNNIHQT 365
QY 338 TLTNFTTSLNLOYSRPMGKGSATFNSTGVLQH-----LRLPLFOKSGMPYLGCOLIS 392
DB 366 TERFRLIPSLLEVKKDVELDELTLW--AADILEHMKSIITLNDPVIATSKIRVLTG---FL 420
QY 393 LRPEKD-----GAATGVDTCYHPDPVPGGLDIOQLYWELSQL--THGVTOLGFVYLD 446
DB 421 LNDIDPFMRVSLTSLTSLATV--DGSNDHODIGQLDFQLERMRRETYLTVEATIHLEGLD 478
QY 447 SLF-----INGYAPONLS-IRG-----EYQINFIIVMNNLSNPPTSSEYITLLRDID 494
DB 479 KIMDAKPLVNGREIMQIALKSGSLIRWQOK--LITWQLAYPNGTAECEKEMKRDYKA 536
QY 495 K 495
DB 537 K 537

RESULT 11

US-11-056-355B-82536

Sequence 82536, Application US/11056355B

Publication No. US20060150283A1

GENERAL INFORMATION:

APPLICANT: Brover, Vyacheslav

APPLICANT: Alexandrov, Nickolai

TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding

TITLE OF INVENTION: Polypeptides Encoded Thereby

FILE REFERENCE: 2750-1590PUS2

CURRENT APPLICATION NUMBER: US/11/056,355B

CURRENT FILING DATE: 2005-02-14

PRIOR APPLICATION NUMBER: 60/544,190

PRIOR FILING DATE: 2004-02-13

NUMBER OF SEQ ID NOS: 119966

SEQ ID NO 82536

LENGTH: 605

TYPE: prt

ORGANISM: Arabidopsis thaliana

FEATURE:

NAME/KEY: peptide

LOCATION: (1)..(605)

OTHER INFORMATION: Ceres Seq. ID no. 12666069

US-11-056-355B-82536

Query Match 3.4%; Score 136; DB 7; Length 605;
Best Local Similarity 20.6%; Pred. No. 0.046;
Matches 99; Conservative 67; Mismatches 185; Indels 130; Gaps 20;

113 KDGATGVDAICTHRPDPGGLDREQLYELSQLTHSITELGPTTLDRDLSLVNGFTHR 172

DB 152 RDEEVOG-DTVIERND-----QSKLETAKL-----RIYDQWIDVNLRESEYEN 197

QY 173 SSVPTTSGVSEEPPTLNTNNLRVYADMGQPSL-KFNITDVMKHLSPFORSS- 230

DB 198 SRIPTMKFGTAKDAPFRDLTINSLFYNNINGAVEDLTERGIDDLKSGKITVPLPAKATF 257

QY 231 -----LGAATGCRVIALRSVNGAETRDLLCTYLQPLSGFGLPIKOVFHEL 278

DB 238 LDDPLRLVLAVERGARFGFTLDEELKEAASSEVERVAL-----GEKISRERIGNEI 308

QY 279 SQQTG-----ITRLGPYSLDKSLYLNGVNEPDPED-----PTTPKP 317

DB 309 DLMISNGPVSATVITLSDKLFSVVPALPSSAEPSPENCGLSOSYLEAMMSLKTTPR 368

QY 318 -----ATFLP-----PLSEATTAMGYHLK 337

DB 369 GKFSGEQRRLALYAAAFPLPRKTVYKDTGKSIPIVNNHIFKFSMKRKTSDAETVNNIHQT 428

QY 338 TLTNFTTSLNLOYSRPMGKGSATFNSTGVLQH-----LRLPLFOKSGMPYLGCOLIS 392

DB 429 TERFRLIPSLLEVKKDVELDELTLW--AADILEHMKSIITLNDPVIATSKIRVLTG---FL 483

QY 393 LRPEKD-----GAATGVDTCYHPDPVPGGLDIOQLYWELSQL--THGVTOLGFVYLD 446

DB 484 LNDIDPFMRVSLTSLTSLATV--DGSNDHODIGQLDFQLERMRRETYLTVEATIHLEGLD 541

QY 447 SLF-----INGYAPONLS-IRG-----EYQINFIIVMNNLSNPPTSSEYITLLRDID 494

DB 542 KIMDAKPLVNGREIMQIALKSGSLIRWQOK--LITWQLAYPNGTAECEKEMKRDYKA 599

QY 495 K 495

DB 600 K 600

RESULT 12

US-11-056-355B-82535

Sequence 82535, Application US/11056355B

Publication No. US20060150283A1

GENERAL INFORMATION:

APPLICANT: Brover, Vyacheslav

APPLICANT: Alexandrov, Nickolai

TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding

TITLE OF INVENTION: Polypeptides Encoded Thereby

FILE REFERENCE: 2750-1590PUS2

CURRENT APPLICATION NUMBER: US/11/056,355B

CURRENT FILING DATE: 2005-02-14

PRIOR APPLICATION NUMBER: 60/544,190

PRIOR FILING DATE: 2004-02-13

NUMBER OF SEQ ID NOS: 119966

SEQ ID NO 82535

LENGTH: 609

TYPE: prt

ORGANISM: Arabidopsis thaliana

FEATURE:

NAME/KEY: peptide

LOCATION: (1)..(609)

OTHER INFORMATION: Ceres Seq. ID no. 12666068

US-11-056-355B-82535

Query Match 3.4%; Score 136; DB 7; Length 609;
Best Local Similarity 20.6%; Pred. No. 0.047;
Matches 99; Conservative 67; Mismatches 185; Indels 130; Gaps 20;

QY 113 KDGATGVDAICTHRPDPGGLDREQLYELSQLTHSITELGPTTLDRDLSLVNGFTHR 172

DB 156 RDEEVOG-DTVIERND-----QSKLETAKL-----RIYDQWIDVNLRESEYEN 201

QY 173 SSVPTTSGVSEEPPTLNTNNLRVYADMGQPSL-KFNITDVMKHLSPFORSS- 230

DB 202 SRIPTMKFGTAKDAPFRDLTINSLFYNNINGAVEDLTERGIDDLKSGKITVPLPAKATF 261


```

QY 231 -----IGARYTGRVIALRSVKNGAETRVLDLCTYLOPLSGPGLPIKQVYHEL 278
DB 262 LDPLRLRAVRRCARGFTLDELKGAASSEVRVAL-----GKXISHERIGNEI 312
QY 279 SQOOTHG-----ITRLGPYSIDKDSLXYLNGVNEGPDEP-----PTTPKP 317
DB 313 DLMISGNGPVSATYVLDLKFVVPALPSSAEPSPENCGLSQSYLAWMSILKTPRP 372
QY 318 -----ATTLP-----PLSEATTMAGYHLK 337
DB 373 GKFSGEORLALYAMELPFKRTYKDTGKKSIPVNHIFKSKRTSDAETVMNHQT 432
QY 338 TLTLNFTISNLQYSPDMKGSATFNSTEGVLQ-----ILRPLFOKSSMGPFYLGCOLIS 392
DB 433 TERFSLPLSLVKKVDELDELTM--AADILEMKSTITLNDPVIATSKIRVLNG---FL 487
QY 393 LRPEKD-----GAATGVDTCTYHDPVPVPGGLDIOQLYWEISQL--THGVTOLGFTVLDRD 446
DB 488 LRDKDFWRVSLTSLISLATV--DGSNDHODIQGLDFQERMETVLTVEATIHGLD 545
QY 447 SLF-----INGYAPQNLIS-ING-----EYQINFIHVMNLSDPPTSEVITTLRDID 494
DB 546 KIMDAKPLVNGREIMQIAELKGSRLIREWQOK--LITWQLAYPVGTABECKEMWRDICA 603
QY 495 K 495
DB 604 K 604

RESULT 13
US-11-123-692-34
; Sequence 34, Application US/11123692
; Publication No. US20060110827A1
; GENERAL INFORMATION:
; APPLICANT: LAGARIAS, JOHN C.
; APPLICANT: FISCHER, AMANDA J.
; TITLE OF INVENTION: A RED AND NEAR INFRARED FLUORESCENT PHYTOCHROMS
; FILE REFERENCE: 407T-303230US
; CURRENT APPLICATION NUMBER: US/11/123,692
; PRIOR FILING DATE: 2005-05-05
; PRIOR APPLICATION NUMBER: US 60/569,310
; PRIOR FILING DATE: 2004-05-06
; PRIOR APPLICATION NUMBER: US60/598,661
; PRIOR FILING DATE: 2004-08-03
; PRIOR APPLICATION NUMBER: US60/640,867
; PRIOR FILING DATE: 2004-12-30
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 34
; LENGTH: 774
; TYPE: PRT
; ORGANISM: Syncechocystis sp.
US-11-123-692-34

Query Match 3.3%; Score 132; DB 7; Length 774;
Best Local Similarity 19.1%; Pred. No. 0.13; Indels 318; Gaps 37;
Matches 155; Conservative 88; Mismatches 252;

QY 173 SSVPSTTGVVSEEPFTL-----NFTINRLRYMADMGQGSIK----- 210
DB 43 SQTASANTGILGRSPEDLGRITLGEVPSFDIDPISRLITAGQSSINPSKLMARVAGDD 102
QY 211 FNITDNVKKHLSPLFORSSISGARYTGCRVIALRSVKNGAETRVLDLCTYLOPLSGPGLP 270
DB 103 FVIDGV-----FRNSDKL-----LVCELEPAVTSNLP 132
QY 271 IKQVFF-----ELSOOTH-----GITRLGPYSLDKS----- 297
DB 133 FLGYFHMANALNLRQOANLRDFYDVIIVEVRMTGFDKVMYLRFDENHGVYIADKR 192
QY 298 -----LYLNGVNEPQDEP-----PTTP--KPATFLPPLSEA 328
DB 193 DDMEPYLGLHYPSDITQPARRLFIHNPRIVIDVGVAVPLPFAVAVPSTNRADVLTES 251

```

```

QY 329 TTMAGYHLKTLTLNFTISNLQYSPDMKGSATFNSTEGVLQHLRLPFOKSSMGPFYL 386
DB 252 ILRSAYH-----CHLTLYLKNMGVAGSLTSLIKOSHLLGLACHHQTKRVPFEL 301
QY 387 --GCOL-----ISLRPEK-----DGAAATGVDTT--CTYHPDE-- 414
DB 302 KRACEFFGRVVFNSISAQETETFDYRVQLAEBHVALDGMTTAAADVEGLTNNPDLIG 361
QY 415 -----VGPGLDIOQLYWEISQLTHGVTOLGFTV-----LDDNSLFTNG 452
DB 362 LTGSGGALICFGEKLLIVGSTEPRDEKAVQYLLQWLEHNEVDVFTSSLSQIYPPAVNFKS 421
QY 453 YAPQNLIS--RGHYQINF-----HYNNMNLSDPPTSEVITTLRDIDQKVTLLYGSQ 504
DB 422 VASGLAIPIARHNFLLMFREPVQLTNV--GQDPNHA-----YEATQ 462
QY 505 -----LHDTFRFCLVTNLT-----MDSVLTVYAL-----FSSNL 534
DB 463 EDGKIBLHPROSFDLMKEIYRLOSPLPMQSVETGALAKKALIVLILRQAEELAQALANL 522
QY 535 DPSLVE-----QVFLDKTLNLSFHWLGSYQLVDIHYTE-----MES 571
DB 523 BRSNADLKFPAYIASHDLQEPINQVSNVYQLLEKRYSEALDEDAKDFIDFAVTGVSILMOT 582
QY 572 SVYQPTSSSTQHYFLNFTTNLPSQDKAPGTNNQRNKNIEDA----- 618
DB 583 LIDDLITVAKVDTQYAOULTFDVQEVVDKALANT-----KORIESSABEIVGSMRAVM 636
QY 619 -----LNLQFRN-----SSIX-----SYFSDCVSTFRSVPNRHNTGVDSLCSFPLA 661
DB 637 ADQIQLMQVFOURLANGIKFPADKSPKIKINGDRQEDAMVAVQONGIGID-----P 688
QY 662 RRVDRVAIYEBELRW--TNGTQLONFLDRSSVLVDGSPN--RNEPLTQNS--ADIQH 715
DB 689 QEFELIFVI--FORLHTRDEYKGTGMGLAICKKLIEGHQGIWLESNPGESSTFYFSIPT 746
QY 716 SGRSSLEGPREFQULISEEDLANMTGHHNNH 748
DB 747 GNGRSFL-----EQKLISEEDLNSAVDRHHNNH 774

RESULT 14
US-11-234-587-355
; Sequence 355, Application US/11234587
; Publication No. US20060194220A1
; GENERAL INFORMATION:
; APPLICANT: Reddy, Ramachandra
; APPLICANT: Gill, Parkash
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DETECTING
; FILE REFERENCE: VASG-P02-001
; CURRENT APPLICATION NUMBER: US/11/234,587
; PRIOR FILING DATE: 2005-09-23
; PRIOR APPLICATION NUMBER: US 60/612,861
; PRIOR FILING DATE: 2004-09-23
; NUMBER OF SEQ ID NOS: 368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 355
; LENGTH: 401
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-234-587-355

Query Match 3.3%; Score 130; DB 7; Length 401;
Best Local Similarity 32.7%; Pred. No. 0.072; Indels 38; Gaps 8;
Matches 50; Conservative 12; Mismatches 53;

QY 632 FSDCOVSTFRSVP-----NRHHTGVDSLCSFPLARR--VDRVAIYEE----- 672
DB 251 FEPNVYTTDREVPRAVDIKVTRSSPSLSLAAVFPAPAGAMLDYEVKTHEKGAQSPSS 310
QY 673 --FLRMTRNGTQLONFLDR--SSVLVD-----GYFP-----RNNEPLTQNSADIQSGG 718

```

```

Db      311  VRFKTSNRAELRG--LKRGAAYLVQVRASEAAGYGPFGQEHHSQTQLDESEGWRQGG 368
QY      719  RSSLEGPRFEQKLISEEDLNM--HTGHHHHH 748
Db      369  RSSLEGPRFEQKLI PNPLGLDSTRGTGHHHHH 401

```

RESULT 15

```

US-11-234-482-15
; Sequence 15, Application US/11234482
; Publication No. US20060204512A1
; GENERAL INFORMATION:
; APPLICANT: Krasnoperov, Valery
; APPLICANT: Zozulya, Sergey
; APPLICANT: Kertesz, Nathalie
; APPLICANT: Reddy, Ramachandra
; APPLICANT: Gili, Parkash
; TITLE OF INVENTION: POLYPEPTIDE COMPOUNDS FOR INHIBITING
; TITLE OF INVENTION: ANGIOGENESIS AND TUMOR GROWTH
; FILE REFERENCE: VASG-P01-004
; CURRENT APPLICATION NUMBER: US/11/234,482
; CURRENT FILING DATE: 2005-09-23
; PRIOR APPLICATION NUMBER: US 60/612,488
; PRIOR FILING DATE: 2004-09-23
; NUMBER OF SEQ. ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 401
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Recombinant B4-CF2
US-11-234-482-15

```

```

Query Match      3.3%; Score 130; DB 7; Length 401;
Best Local Similarity 32.7%; Pred. No. 0.072;
Matches 50; Conservative 12; Mismatches 53; Indels 38; Gaps 8;

```

```

QY      632  FSDCQVSTFRSVP-----NRHRTGVDSL CNFSPLARR--VDRVAIYER----- 672
Db      251  FEPVNTTDTREVP PAVSDIRVTRSSPSSLSLAWAVPRAPSGAWLDYEVKYEKGAEGPSS 310
QY      673  --FLRMTNRNGTQLONFILDR-SSVLYD-----GYSF---NRNEPLTGN SADIQHSGG 718
Db      311  VRFLKTSNRAELRG--LKRGAAYLVQVRASEAAGYGPFGQEHHSQTQLDESEGWRQGG 368
QY      719  RSSLEGPRFEQKLISEEDLNM--HTGHHHHH 748
Db      369  RSSLEGPRFEQKLI PNPLGLDSTRGTGHHHHH 401

```

Search completed: October 14, 2006, 04:10:05
 Job time : 43 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: October 14, 2006, 03:59:20 ; Search time 46 Seconds
(without alignments)
1564.568 Million cell updates/sec

Title: US-10-687-035-1

Perfect score: 3945

Sequence: 1 AAQPARARARLFTFRSSV.....OKLISEDLNMTGHHHHH 748

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Listing first 45 summaries

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|----------|----------------------|
| 1 | 144.5 | 3.7 | 1408 | 2 H69068 | cell surface glyco |
| 2 | 135.5 | 3.4 | 1117 | 2 S63399 | probable membrane |
| 3 | 123.5 | 3.1 | 2477 | 2 S14428 | fibronectin precu |
| 4 | 122 | 3.1 | 631 | 2 I52257 | episialin - mouse |
| 5 | 121.5 | 3.1 | 1475 | 2 S42718 | nuclear pore compl |
| 6 | 121.5 | 3.1 | 2471 | 2 T42977 | large tegument pro |
| 7 | 121 | 3.1 | 5005 | 2 F82884 | hypothetical prote |
| 8 | 121 | 3.1 | 1131 | 2 T41144 | hypothetical serin |
| 9 | 118 | 3.0 | 532 | 2 T34235 | hypothetical prote |
| 10 | 117.5 | 3.0 | 1377 | 2 A38926 | DNA-binding protei |
| 11 | 116 | 2.9 | 1848 | 2 A44140 | cellulose-binding |
| 12 | 115.5 | 2.9 | 1474 | 2 F69009 | probable membrane |
| 13 | 115 | 2.9 | 888 | 2 A54280 | cell differentiation |
| 14 | 114.5 | 2.9 | 709 | 2 A35364 | carcinoembryonic a |
| 15 | 113.5 | 2.9 | 855 | 2 T07015 | Cf-4A protein - to |
| 16 | 113.5 | 2.9 | 1004 | 2 B25039 | outer cell wall pr |
| 17 | 113.5 | 2.9 | 1369 | 2 T17504 | hypothetical prote |
| 18 | 113.5 | 2.9 | 3344 | 2 UQ1899 | genome polypeptid |
| 19 | 113.5 | 2.9 | 5762 | 2 A41819 | proline-rich pepti |
| 20 | 112.5 | 2.9 | 838 | 2 S38156 | hypothetical prote |
| 21 | 112.5 | 2.9 | 948 | 2 T11678 | hypothetical prote |
| 22 | 112.5 | 2.9 | 1348 | 2 AH1115 | cell surface prote |
| 23 | 112.5 | 2.9 | 4936 | 2 AH2515 | hypothetical prote |
| 24 | 112 | 2.8 | 643 | 2 S55593 | membrane protein S |
| 25 | 112 | 2.8 | 3587 | 2 T31075 | tyrocidine synthe |
| 26 | 111.5 | 2.8 | 919 | 2 S45889 | probable regulator |
| 27 | 111.5 | 2.8 | 1622 | 2 AE1717 | probable cell surf |
| 28 | 111 | 2.8 | 1777 | 2 T34369 | hypothetical prote |
| 29 | 111 | 2.8 | 2386 | 1 FNHU | fibronectin precu |

| | | | | | |
|----|-------|-----|------|----------|---------------------|
| 30 | 110.5 | 2.8 | 630 | 2 A39344 | tumor-associated m |
| 31 | 110.5 | 2.8 | 1778 | 2 J70382 | apolipoprotein B - |
| 32 | 110.5 | 2.8 | 2508 | 2 S61441 | surface-associated |
| 33 | 110 | 2.8 | 628 | 1 A39262 | transcription fact |
| 34 | 110 | 2.8 | 1128 | 2 A49960 | bud emergence prote |
| 35 | 110 | 2.8 | 1349 | 2 A11476 | cell surface prote |
| 36 | 110 | 2.8 | 3131 | 2 T39553 | vacuolar protein s |
| 37 | 109.5 | 2.8 | 1268 | 1 A39640 | neural cell adhe |
| 38 | 109 | 2.8 | 989 | 2 T56333 | apolipoprotein B - |
| 39 | 109 | 2.8 | 998 | 2 C75489 | conserved hypothet |
| 40 | 108.5 | 2.8 | 500 | 2 D97302 | hypothetical prote |
| 41 | 108.5 | 2.8 | 504 | 2 T33485 | hypothetical prote |
| 42 | 108.5 | 2.8 | 537 | 2 A35400 | surface protein T6 |
| 43 | 108.5 | 2.8 | 776 | 2 A82787 | TonB-dependent rec |
| 44 | 108.5 | 2.8 | 1040 | 2 D88568 | protein ZK757.3 (1 |
| 45 | 108.5 | 2.8 | 1113 | 2 H84105 | hypothetical |

ALIGNMENTS

RESULT 1

H69068 cell surface glycoprotein (8-layer protein) related protein - Methanobacterium thermoaut
C:Species: Methanobacterium thermoautotrophicum
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C:Accession: H69068
R:Smith, D.R.; Doncette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.
K.; S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: func
A:Reference number: A69000; PMID:98037514; PMID:9371463
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Accession: H69068
A:Molecule type: DNA
A:Residues: 1-1408 <MTH>
A:Cross-References: UNIPROT:O27557; UNIPARC:UPI000006761; GB:AE000911; GB:AE000666; NID
A:Experimental source: strain Delta H
C:Genetics:
A:Gene: MTH1513
A:Start codon: TTG

| Query Match | 3.7% | Score 144.5 | DB 2 | Length 1408 | Best Local Similarity 20.1% | Pred. No. 0.19 | Matches 157 | Conservative 99 | Mismatches 258 | Indels 269 | Gaps 40 |
|-------------|------|--|----------------------------------|-------------|-----------------------------|----------------|-------------|-----------------|----------------|------------|---------|
| QY | 58 | NETITNLRYEENMMPGSKKFTT--- | ERYVGLRLPLFKNTSVGPVYSGCHTLTPREKD | 114 | | | | | | | |
| DB | 444 | NFTVY-----PDACKKVLTLGSDGLTRFTPTIISNLSVPL----- | | 480 | | | | | | | |
| QY | 115 | GEATGVDAICTHRPPPGGLDRQLYLEL-SQTHSTTEGPTLDRDSLTVNGFTARS | 173 | | | | | | | | |
| DB | 481 | -----NGICP-----LVYTLRGNTITNGDSTGWRDL--LYNGY--RT | 515 | | | | | | | | |
| QY | 174 | SVPTTSGVSEEPFTLNFTI----- | NNLRVYADNGQ | 205 | | | | | | | |
| DB | 516 | AGKWEVGSLETPEFEFEYTLRNSRAVSVGNPPATVRFGLPLIPENILRVTPSGGQ | 575 | | | | | | | | |
| QY | 206 | PGSLKFNITDVMGHLSPLQFSSLGARYGCIVALRSYKNAERFV----- | 254 | | | | | | | | |
| DB | 576 | E-PLKVVNTADLVVYGDLPSDYTEL--YIDGVLLDSRNTVNAASSFTTVSFNRTLAAG | 631 | | | | | | | | |
| QY | 255 | --DLICTVLP-----LSGFLPIKOVFHELSQOHTG-----ITRL--GPSYSDK | 295 | | | | | | | | |
| DB | 632 | LYETINDLELYVWEGEKFTYEN--FTITPGGAAPLTVVYSAMTINDSNPRSYTA | 689 | | | | | | | | |
| QY | 296 | DSLTVNG-----YNEPGDEPPTPKPATFPLPLPSEATVANGYH---LKTULT-LNF | 343 | | | | | | | | |
| DB | 690 | -TIVNGVPHTKVLTNIPGES---TVPESTISLLPDGGLYTIISNNVSGTVRVLSANF | 745 | | | | | | | | |
| QY | 344 | TISNLOYSPPMKGSALFNSTEGVLQHLRLPLFKSSMGPRVYLCQLSLRPEKGATG | 403 | | | | | | | | |

```

Db      746 TLSNVTSPVEKSPINVTVAIV-----RNGDLAG 777
Qy      404 VDTTCYHPDPVGPGLDIQQLYWEISQLTGVTQLGFFYLDRLSLINGAAPONLSIRGE 463
Db      778 -DAVTLVLDVA-----METRTVS-----VPGKSVLV--FKELAPGE 816
Qy      464 YQINFIH---VNMNLSNPDPTSSSEYITLRLDIDQKVTTLKGSQLDHTFACL-VTN--- 516
Db      817 YRLNLSGTDTVTTRVLEPDLITIGF-----NTPVTPAPL--SVASLANTNPHD 865
Qy      517 -----LTMDSVLYTKALFSSNLDPSLVEQVFLDKTNASFWMLG----- 566
Db      866 LVIGETARLWDDGVVENIV--SLSPGETREIAMGTLTPGHTVGINEFSKIVAVLR 922
Qy      557 -STYQVLDIHYTEMESVYOP-----TSSSTQHFYLNFT---TNLPISQKAPGT 605
Db      923 PASITLSDLRVT--PSSGFSPLITATATANTGEVDGNVAVLYINGLAVDERKNTVGA 980
Qy      606 TNYQRNKNIEDALNOLFRNSI-----KSFSDCOVSTRSVF---NRHHT 649
Db      981 -----GRSVQVAFNHTIENAGIYLAGISLTPLDVRLSEPAISNLATPLTGVSPHRI 1034
Qy      650 GVDSLCNS-----PLARRVRYAIEEFLRMTRNGTQLQNTLDRS-----SVLVDS 697
Db      1035 IVTALVSTTERSGSNYTAGLYIDGVNVQNRVRYVTGPSVLVSFTADISEPGEHQVTNS 1094
Qy      698 YSP 700
Db      1095 LSP 1097

```

RESULT 2

663399
 probable membrane protein YNR067C - Yeast (Saccharomyces cerevisiae)
 N/A: alternate names: hypothetical protein N3547
 C/Species: Saccharomyces cerevisiae
 C/Date: 27-Apr-1996 #sequence_revision 03-May-1996 #text_change 09-Jul-2004
 C/Accession: S63399
 R/Duesterhoeft, A.; Floeth, M.; Fritz, C.; Heuss-Neitzel, D.; Hilbert, H.; Moestl, D.
 submitted to the Protein Sequence Database, April 1996
 A/Reference number: S62944
 A/Accession: S63399
 A/Molecule type: DNA
 A/Residues: 1-1117 <DURS>
 A/Cross-references: UNIPROT:P53753; UNIPARC:UPI000013BACB; EMBL:Z71682; NID:G1302597; P1
 A/Experimental source: strain S288C
 C/Genetics:
 A/Genes: MIPS:YNR067C
 A/Cross-references: SGD:S0005350
 A/Map position: 14R
 C/Keywords: transmembrane protein
 F/6-22/Domain: transmembrane #status predicted <TM>

```

Query Match      3.44; Score 135.5; DB 2; Length 1117;
Best local similarity 21.1%; Pred. No. 0.56; 255; Indels 233; Gaps 39;
Matches 156; Conservative 94; Mismatches 255;
Qy      4 PARARATKLFTHRSVSTSTPCTPTVYIGA-----SKTPAS--IFGSAASHL 51
Db      206 PSYSSQETKIIP--SLTSSNKTITITISVRTAAATAGDSFIACSTPSSSTLFYPSNSTOP 263
Qy      52 LIIFLFTTNLTYEEN-----MPPGSRKNTTERTV-----LOGLLRPL 91
Db      264 LVQ-TLASTSTSPAVPSNRQTITLSPSVLSYTSIPSPNITENGSPSPSLSTVSPV 322
Qy      92 FKNTSVGPVLYSGCRITLIRPKQGEATVNDALICHRPPPTGRLDRLQYELG-QLTHS 150
Db      323 YPSSSTGNIL-----LSSLFSTVDSSS-----PVSGTLD--TIYSSSMQATIS 365
Qy      151 ITELQPYTLUDSLVYNFTHRSSVPTSTGVSEFPITLFTTNLRYMADMQPGSLK 210
Db      366 SSSSRQRTKSSS-----SLSTSTSTATTNTSSST--TIYNLNAVSTDEPPTV- 414

```

```

Qy      211 FNITDWMKHLSPLEFRSSSLGARFTGCVIALRSVKGAEFRVDLCTYLQPLSGPGPL 270
Db      415 FDRSPNPM-----SLADEVSN-----DGP-IG 435
Qy      271 IKQVFEHL---SQOTHGTRLGPYSLDK-----DSLLYNGNEPEGDEEP 312
Db      436 TNKFYTNLVGSGESAPF--VYPYSLMKTSSSYGFAVQHTTVDDQSYSGYSGSAGAEVL 493
Qy      313 TTPKPATFLPPLSEATYAMGHYKLTLTNFTSNLOYPDMQKGSATFNSTEGVL----- 368
Db      494 VNP-----LGIAHVFSASNPDSSMTWQVDENITLSTRTVLSSEN 533
Qy      369 ---QHLLRPLFKSSMGPFLYGQLSLRPEKGAATGVDTCTYHPDPVGPGLDIQQLY 425
Db      534 DSSNYLEIPLVQ--GNG-FATGTYHSLN-AKIGSSVGFITIVSESSNLAQGI-----LK 585
Qy      426 WEISQLTGVTQCGFYLDRLSLFINGYAPONLSIRGEYOINH-----IYNMNLSNP 479
Db      586 YRIT-LINGVTWL-CYVIGPDDLSTDF---SLEVSSEYEIKASAVDGHILQAVN-PS 639
Qy      480 PTSGE-----YIT-LLRDIQKVTTLTK-----GQOL-----HDTFR 510
Db      640 ETDYEVFYDQAGMYVTNPFILQGVSDSTAYTEFSYTTQGESASGSTMIRALPHHESFS 699
Qy      511 FCLVTLTMSVLTVAKALFSSNLDPSLVEQVFLDKTNASFWMLGSTYQ----- 560
Db      700 DIMQDYVTGILASTTGVMNGYLTSLQ---FSTSLNQISLWLPSSQLGSNLEYSK 755
Qy      561 -----LYDIIHYTEMESVYQPTSSSTQHFYLNFTTNLPY-----GDKAKAPGTN 607
Db      756 EQLQDLAEVANSSEIQVSSISGLNT--YLVGVIDKYSYLLTVSEIILQDEASTYST- 812
Qy      608 YQRNKNIEDALNOLFRN 625
Db      813 ----LENIKSAFDILQN 826

```

RESULT 3

S14428
 fibronectin precursor - rat
 C/Species: Rattus norvegicus (Norway rat)
 C/Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 09-Jul-2004
 C/Accession: S14428; S12455; A22319; S46203; S00459; A27252; I59049
 R/Hynes, R.O.
 submitted to the EMBL Data Library, July 1989
 A/Reference number: S14428
 A/Accession: S14428
 A/Molecule type: mRNA
 A/Residues: 1-2477 <HN>
 A/Cross-references: UNIPROT:P04937; UNIPARC:UPI000012A7C6; EMBL:X15906; P1DN
 R/Schwarzbauer, J.E.; Patel, R.S.; Fonda, D.; Hynes, R.O.
 EMO J. 6, 2573-2580, 1987
 A/Title: Multiple sites of alternative splicing of the rat fibronectin gene transcript.
 A/Reference number: S12455; MUID:88054951; PMID:2445560
 A/Accession: S12455
 A/Molecule type: mRNA
 A/Status: nucleic acid sequence not shown
 A/Residues: 609-1810, 'T', 1812-2283 <SCH>
 A/Cross-references: UNIPARC:UPI0000177AEB; EMBL:X15906
 R/Tamkun, J.W.; Schwarzbauer, J.E.; Hynes, R.O.
 Proc. Natl. Acad. Sci. U.S.A. 81, 5140-5144, 1984
 A/Title: A single rat fibronectin gene generates three different mRNAs by alternative splicing.
 A/Accession: A22319; MUID:84238097; PMID:6089177
 A/Molecule type: DNA
 A/Residues: 2052-2237 <TM>
 A/Cross-references: UNIPARC:UPI0000177AEB
 R/Falkenberg, C.; English, J.J.; Thøgersen, I.B.; Salvesen, G.; Akerstrøm, B.
 Biochem. J. 301, 745-751, 1994
 A/Title: Isolation and characterization of fibronectin-alpha(1)-microglobulin complex 1
 A/Reference number: S46203; MUID:94330948; PMID:7519849
 A/Accession: S46203
 A/Status: preliminary

A:Molecule type: protein
A:Residues: 1163-1192/'GAIN',1268,'P',1270-1271,'D',1273,'CR',1276,'PY',1385-1399 <FAL>
A:Cross-references: UNIPARC:PIP0000177AEC; UNIPARC:PIP0000177AED; UNIPARC:PIP0000177AEE
R:Patel, R.S.; Odermatt, E.; Schwarzbauer, J.E.; Hynes, R.O.
EMBO J. 6, 2565-2572, 1987
A:Title: Organization of the fibronectin gene provides evidence for exon shuffling during
A:Reference number: S00459, MUID:88054950, PMID:3119323
A:Accession: S00459
A:Molecule type: DNA
A:Residues: 1-139/'2382-2477 <PAT>
A:Cross-references: UNIPARC:PIP000017095C; UNIPARC:PIP0000177AEF; EMBL:X05631
A:Note: The authors translated the codon CCG for residues 51 and 94 as Ala
R:Schwarzbauer, J.E.; Tamkun, J.W.; Lemischka, I.R.; Hynes, R.O.
Cell 35, 421-431, 1983
A:Title: Three different fibronectin mRNAs arise by alternative splicing within the cod
A:Reference number: A27252; MUID:84082067; PMID:6317187
A:Accession: A27252
A:Molecule type: mRNA
A:Residues: 1586-1720,'T',1722,1813-2477 <SC2>
A:Cross-references: UNIPARC:UPI0000177AF0
R:Odermatt, E.; Tamkun, J.W.; Hynes, R.O.
Proc. Natl. Acad. Sci. U.S.A. 82, 6571-6575, 1985
A:Title: Repeating modular structure of the fibronectin gene: Relationship to protein st
A:Reference number: I59049, MUID:86016741; PMID:3863113
A:Accession: I59049
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1722-1810 <RES>
A:Cross-references: UNIPARC:UPI00000040C; GB:M11750; NID:G204164; PIDN:AAA41170.1; PID
C:Genetics:
A:Anticodon: 51/1; 94/1; 2416/3; 2454/3
C:Superfamily: fibronectin; fibronectin type I repeat homology; fibronectin type II rep
C:Keywords: alternative splicing; cell adhesion; collagen binding; disulfide bond; dupli
F:1-32/Domain: signal sequence #status predicted <SIG>
F:33-2477/Product: fibronectin #status predicted <MAT>
F:53-88/Domain: fibronectin type I repeat homology <1F1>
F:98-136/Domain: fibronectin type I repeat homology <1F2>
F:147-180/Domain: fibronectin type I repeat homology <1F3>
F:187-228/Domain: fibronectin type I repeat homology <1F4>
F:208-271/Domain: fibronectin type I repeat homology <1F5>
F:308-342/Domain: fibronectin type I repeat homology <1F6>
F:360-401/Domain: fibronectin type II repeat homology <2F1>
F:420-461/Domain: fibronectin type II repeat homology <2F2>
F:470-508/Domain: fibronectin type I repeat homology <1F7>
F:518-555/Domain: fibronectin type I repeat homology <1F8>
F:561-599/Domain: fibronectin type I repeat homology <1F9>
F:609-692/Domain: fibronectin type III repeat homology <FN3A>
F:718-800/Domain: fibronectin type III repeat homology <FN3B>
F:809-890/Domain: fibronectin type III repeat homology <FN3C>
F:905-987/Domain: fibronectin type III repeat homology <FN3D>
F:995-1076/Domain: fibronectin type III repeat homology <FN3E>
F:1085-1164/Domain: fibronectin type III repeat homology <FN3F>
F:1172-1257/Domain: fibronectin type III repeat homology <FN3G>
F:1265-1348/Domain: fibronectin type III repeat homology <FN3H>
F:1356-1439/Domain: fibronectin type III repeat homology <FN3I>
F:1447-1529/Domain: fibronectin type III repeat homology <FN3J>
F:1537-1619/Domain: fibronectin type III repeat homology <FN3K>
F:1614-1616/Region: cell attachment (R-G-D) motif
F:1631-1713/Domain: fibronectin type III repeat homology <FN3L>
F:1721-1803/Domain: fibronectin type III repeat homology <FN3M>
F:1811-1893/Domain: fibronectin type III repeat homology <FN3N>
F:1903-1984/Domain: fibronectin type III repeat homology <FN3O>
F:1992-2074/Domain: fibronectin type III repeat homology <FN3P>
F:2181-2183/Region: cell attachment (R-G-D) motif
F:2193-2273/Domain: fibronectin type III repeat homology <FN3Q>
F:2296-2335/Domain: fibronectin type I repeat homology <1F10>
F:2341-2378/Domain: fibronectin type I repeat homology <1F11>
F:2385-2420/Domain: fibronectin type I repeat homology <1F12>
F:53-79, 77-88, 98-126, 124-136, 142-170, 168-180, 187-216, 214-226, 232-261, 259-271, 308-335, 333-368, 2366-2378, 2385-2411, 2409-2420/Disulfide bond: #status predicted
F:2458/Disulfide bond: interchain (to 2462) #status predicted
F:2462/Disulfide bond: interchain (to 2458) #status predicted

```

Query March 3.1%; Score 123.5; DB 2; Length 24717;
Best Local Similarity 19.5%; Pred. No: 14;
Matches 177; Conservative 133; Mismatches 309; Indels 287; Gaps 53;

QY 14 FTHRSSVSTVS-----TPTPTVYLGASTKPAISGPSASHLLILFTLNFITNLYRY 66
DB 697 FTTASFTPTVINSYNTGTGTAPEFSPV-----ATSSVTEIRATSSVSVWMSADIVSGERV 751
QY 67 E-----ENMPPGSRKENTERVLOGLRPLPEKNTSVGLYS 102
DB 752 EYELSEEGDEPQYLDLPSTATSVNI PLLLRKRYIVANVYQLSEEGKSLILSTS----- 805
QY 103 GCRLLTLRPEKDGATGVDAICTHRPPPTGEGDLREDQYLELSQTHSITELGTYTLD 162
DB 806 -----QTTAPDA-----PBPPTVDQVDDTISIVRWSRPAAPIT---GYRI-- 842
QY 163 SLY---VNGFTHRSSVPTTSNGVVSSEPFNLFINNLRYMADMGQGSLSKFNITDVMK 219
DB 843 -VSPSPESGGSSTEINLEPTANSV-----TLSDL-----QIG-VQYNITTYAVE 883
QY 220 --HLISPLF--GRSSIG-----ARYTGCRIYALRSVKNGAET--RVDLLCT 259
DB 884 ENGSTPTFIIQGETTGVPRSDVDPAKDLQEVETDVKVTIMTTPRNSAVTGVKRVDLVP 943
QY 260 YLQPLSGEGLPI-KQVTHELSQOHTGIT---RLQPYSLDKDSLTYLNGYNEGPPEP--- 311
DB 944 NLPEHGQGRLEPVNNNTAEVYTGSLSPGTYLFKFAVHQGRESKXPLTQOQTKLDAPTNLQ 1003
QY 312 ---FTTPRPATFPLP-----LSBATAMG---YHLKTLNLFTINSLOVSPMG- 355
DB 1004 FVNEDRTVLVTMPRPARIAGKYLTVGLTRGGQPKQINQPMASKYPLRLQGSSEYTV 1063
QY 356 ----KGSATNSTEGVLOHLRLPEFKSSMGPP-----YLGCQILSL 393
DB 1064 TLMVAKNGQSPKATGVFT--TLQPL---RSIPRPNTEVTETIYITWPAIRIGFKL-GV 1118
QY 394 RPEKDGATGVDTTCTYHPDVGSGDLIDQQLYMWLSQTHGVTLGFTVL-----DRD 446
DB 1119 RPSQGEAPREVT-----DSGSIV--VSGLTGVEYV--YTIQVLRDQGERD 1162
QY 447 SLFINGTA-----PQLSIRGEYQINHIYNNMNS-NPD-----PSSEYITLLRD 491
DB 1163 APIYNRVVTPLSPPTNLHLEANPDGVLTVSWERSTPDTGYRITTTPTNGQGTALBE 1222
QY 492 IQDKVTTLYKKSOLDHFRFCLVTN-----LTMDSVLTVK-----ALFSSNLV---DPSLV 539
DB 1223 V-----VHAQSSCTFENRNPGLERYNSVYTVVADKDSAPISDVIPEVQOLT 1270
QY 540 EQVFLDKT-LNASFHW--LGSYQYLDVHITEMSS-----YQPTSSSSTOHFYALNFTIT 592
DB 1271 DLSFVDLIDSSIGLRMTPIPLNS--TIIGRITVVAAGGIRPIFEDFVDSVGY-----YTVT 1325
QY 593 NLPTFSQDKAQGCTTNYQANKNT-----EDALNQLFRMSST-----KSYFSCQVSTF 640
DB 1326 GL-----EPG-IDYDISVTTLINGESAPITTLQOTAVAPPDRLRFTNIGPRTMVTW 1377
QY 641 RSVENRHHHTGDSLCSNPSPLARVDRAVAYEEPLRMTRNGTOLN-----FTLDRSSVL 694
DB 1378 APPSIEITLN--LVRSVPVKNED--VVELISPDNNAVLTNLLPGTEIYVSVSV- 1431
QY 695 VDGYSPNNNEBLTGNASADIOHSGGRSSIEGPR-PEOKLISEEDLNMH-----TGH-- 743
DB 1432 ---YEQHSISPIRLRQ-----KTGLDSPGTGDSDDVANSPTVWVAPRABITGYII 1480
QY 744 -HHHHH 748
DB 1481 RHHAEH 1486

```


QY 120 -----VDAICT-HRPDPTGPGLDREOLYLELSQLTHTSITELEPGYTLDRDSLIV 166
DB 377 VSSKLKEDEDEGWIDIDPTLVNVSPTTEATNSDQETIYIGDENIHDMWSVDDDT--DDTLDI 434
QY 167 NGFTHRSSVPTTSTGVSEEPFLANFTINNLRVYAMMGOPGSLKFNITDVMGH--LISF 224
DB 435 N-FVQDVKYITSLQNIPIINNTFTIIDKPSNRHKGKMLAIDRLITNIIIEHGLITS 493
QY 225 LFGRSSLGARYTCGRVIALRSVKNGABT-----RVDLCYVQLPSGEGPLIKOVFHE 277
DB 494 ---PNSMSKCKSLQFVFLWSEKLSIPTKDLTKTLTKTLNLTIEIAKVASFKL--TNNIF-- 547
QY 278 LSGQTHGTRTGLGYSLDKSLYLNGVNEPDEPPTPKATFTFLPPLSATATAMKYHLK 337
DB 548 ---KNNITTKLTK--CHEKIKL-----ETQDN-----YKOLLALISKSTTITQYAT 569
QY 338 TLTNFTLISNLQYSPMGKSGATFNSTEG-----VLQHLRLPLF--QKSGMPFYLGC 388
DB 590 EIEELK-NISGM-FTSELGDFSVICTNEBYTALLAIEMLKEKIFGRKQELHAEIYFQS 647
QY 389 QLTISLPEKDGAAATGVDTCTYHPDPVPG-LDIQQLYWELSQLTHGVTQLGFYVLDKRS 447
DB 648 VTIAME-----TFQPIPLPTKYLEIQ-----LEKSK 674
QY 448 LPIINGAPQNLSTIRGEYQIN-----FHYVMNLSPNDPTSSXYITLLRDIOKATTVLK 501
DB 675 VFLEKLKPEQKLTTE--VNEELLTDLHLNKQDATERLPV-PDFTTLKNIQSTLLDLH- 730
QY 502 GSQLHDTFRCLVTNLTMDSVLVTKALFFSSNDPSLVEQFLDKTLNASFHWLGSTYOL 561
DB 731 -----TCYTDLANIDKEVI-----GSTIQ- 748
QY 562 VDIHVTMESSYQYQPTSSSTQHFYINFTITNLPSQDKAQPGT-----TNYQ 609
DB 749 -----QLSYIGWEVAELSHQWNPSPASPIIPLQLDELKTEVQ 787
QY 610 R--NRKNIEDALNQLFRNSSIKSYFSDQVSTFRSPNNRH-----TGVDSLCNFSP 659
DB 788 RYAKQOCTETLNOIL--SPVOELLSQAEQSSSTLSIPILQHYITQAGTVLGESEKNETFES 845
QY 660 LARRVDVAIYEELFMTRNGTOLQNTL--DRSSVLDVGYSPNNREPL 706
DB 846 LRPDTVQKLSIEEFLKMLNSTLTLLNIQIQIKDISLISNQYMOSETI 895

RESULT 7
F82884
hypothetical protein UH495 (imported) - Ureaplasma urealyticum
C/Spectes: Ureaplasma urealyticum
C/Date: 18-Aug-2000 #sequence_rev1510n 20-Aug-2000 #ext_change 20-Aug-2000
C/Accession: F82884
R:Glaas, J.I.; Lefkowitz, E.J.; Glaas, J.S.; Heiner, C.R.; Chen, E.Y.; Caswell, G.H.
submitted to GenBank, February 2000
A/Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a mit
A/Reference number: A82870
A/Accession: F82884
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-5005 <Gla>
A/Cross-references: UNIPARC:UPI0000110214; GB:AE002148; GB:AF222894; NID:G6899495; PIDN:
A/Experimental source: serovar 3; biovar 1
C/Genetics:
A/Gene: UH495
A/Genetic code: SGC3

Query Match 3.1%; Score 121.5; DB 2; Length 5005;
Best Local Similarity 19.6%; Pred. No. 58;
Matches 135; Conservative 108; Mismatches 216; Indels 229; Gaps 39;
QY 136 DREOLYLELSQLTHTSITELEPGYTLDRDSLIVNGFTHRSSVPTTSTGVSEEPFLANFTI 194
DB 692 NHEQVY-ELANSPQSLSEFG-----NELLVNTYINPEPNTITDYKTVLANKVAKLTILKRS 744

QY 195 N-----NLRYAMDG-----QPSL-----KENITD-NVMKHLSPLFORSISIGAR 234
DB 745 NLMEYELKLTITIDNNGDVVWSDPVSINNVMNEYENLDDKALK-----SRIYK 795
QY 235 YTCGRVIA-LRVKNGAETRVLDLCYVQLPSGGLPIKOVFHELSQTHGTRTGLGYPYST 293
DB 796 FGLVYFKQVNVNITKSANKVM-----NNTIPLOIQTA 829
QY 294 DKOSLYLNGVNEPDEPPTPKATFTFLPPLS-----EATLMGTHLKTLLTN 342
DB 830 SKINL-----NSPKVTINNITPASADLYPLVSTDDIPAKDOVIDVTIGNKNTKN-DLM 863
QY 343 FTISNLQYSPMGKSGATFNSTEGVQLH--LRP-----LFQKSGMPFYLGCQL 390
DB 884 FT-SNLEVD-DITK-----TWGAKIHANNLSPEVNYQIKSVKFRQKFLGAFVNV-- 931
QY 391 ISLRPEKDGAAATGVDTCTYHPDPVPGGLDIOQLYWELSQLT-----HGVTOQLGFYVL 443
DB 932 -----NSDNNHVLIDSQKT-----PGSTPKASFDLVSVTASDVVDANVTNSVHIV 979
QY 444 DRDSLIFNGVAPQNLSTIRGEYQI--NPHIVMNLSPND----- 479
DB 980 NNDGTSINACKAKIYNDGEHDLVSNPFVLNVGVNEYDFNLDDLKGNRIYGFKKLIYTE 1039
QY 480 PISSEY-ITLLDIDQKTYTLKSGQLHDTFRCLVTNLTMDSVLVTKALFSSNL----- 534
DB 1040 PQONDYKFSFLNN--QKITNFKTASVNE-----QLTIDTKLVKRP--DSNNKIN 1086
QY 535 -----DP-----SLVEQFLDKTLNASFHWLG-----STYQVLDIHTVMESSYQY 575
DB 1087 LKQLQNDPNDPFLQDSILEITHEDEKDKTHNAVIGKINVDANNKTLSESVENTTFKIQ 1146
QY 576 PTSSSTQHFYINFTITNLPSY-SQDKAQGCT--NYQRKNIEDALNQLFRNSSIKSYRS 633
DB 1147 P-----NHKYL--VDNINATKMKIQPANNISNNSKNSIYDAS--NPSKILSFYN 1194
QY 634 DCQVSTFR-SVPRRHHTGVDSLGNFPLAR-----RVDKVAIYEELFMTRNGTOLQNT 687
DB 1195 EEFVNNINNVQNP-----TNLNPASIDVELKSONLLKDYLRALYIDNNHQKIW 1245
QY 688 LDRSSVLDVGYSPNNREPLTGNADIQH 715
DB 1246 SDYASV-----NNVDLAH 1258

RESULT 8
T41144
hypothetical serine rich protein - fission yeast (Schizosaccharomyces pombe)
C/Spectes: Schizosaccharomyces pombe
C/Date: 03-Dec-1999 #sequence_rev1510n 03-Dec-1999 #ext_change 09-Jul-2004
C/Accession: T41144; T41591
R:Hilbert, H.; Dueserhoef, A.; Wood, V.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, October 1998
A/Reference number: Z21973
A/Accession: T41144
A/Status: translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-1131 <Hil>
A/Cross-references: UNIPROT:O74851; UNIPARC:UPI000006A713; EMBL:AL031907; NID:G3766362;
A/Experimental source: strain 972h-; cosmid c18
R:Murphy, L.; Harris, D.; Lyne, M.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, September 1998
A/Reference number: Z21968
A/Accession: T41591
A/Status: translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 44-1131 <Mur>
A/Cross-references: UNIPARC:UPI0000161F34; EMBL:AL031543; NID:G4239667; PIDN:CAA20837.1
A/Experimental source: strain 972h-; cosmid c74
C/Comment: For a possible alternative initiator product, see PIR:T42367.
C/Genetics:
A/Gene: SPC74.07c; SPDB:SPCC18.01c

A:Map position: 3
C:Keywords: alternative initiators

Query Match 3.1%; Score 121; DB 2; Length 1131;

Best Local Similarity 18.6%; Pred. No. 5.7;
Matches 161; Conservative 130; Mismatches 353; Indels 222; Gaps 30;

```

QY 2 ACPARRARTKLFTRSSVSTSTPCTPVYLGASKTSPASIGPSAASHLLFLFTNFT 61
DB |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 341 AASSSSASATATSSASSIAT--SPTSSSNVSSISTSSMSSAVSSISVQSSIASII 398
DB |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 62 TNLRYEENMPGSRKFTTERVLOGLRLPLFKNTSVGPVSGCRLTLARPEKGEATVD 121
DB |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 122 ---AICTRPDPTRGRLDEQLYELISQTHSITEGLPTLDRDLSLVNGFTHRSSVPT 178
DB |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 446 ASLAVPTVSSSTTG-----SLHYK-TTTTWTVEFTRYLDDSTPPTS---SSIFST 494
DB |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 179 ST-----GVSEBPFTLNFTINLRMYADMQPSLKNITDNVKKHLSPFLQR 228
DB |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 495 ATEADTVSQTSAIYDSSSTSNIGSSSVYASSTGALSNSLSSSTSVSTYIPNAS 554
DB |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 229 SSIGARYTGCRYTALRSYKGAETRVLLCTYLQPLSGPLIKQYFHELSCQTHGTR 288
DB |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 555 SVYASSTALSSNSLSSSTSSAST-----SVIPASSSYEVAASNSDYVQTVSITAS 608
DB |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 289 GPYSLDKDSL-----YLANG-----YNEGPPEPTTPKPAITPLPLSEATTMG 333
DB |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 609 GTTSTSTSEIVSTPASNSNTGSLNGTSSFVNVSGVPSQCTPTPTSSSITGSOLKETSP 668
DB |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 334 YHLKTLTLNFTISNLQYSPDMKGATENS-----TEGVLOH 370
DB |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 669 AYVSS-TVSYTSSVD-----SSSTVNSTGSSSDSGSFGCTTYSDPTTTITSVSGSI 720
DB |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 371 LRLPLFOKSMGPFYLCQLISLRPEKGAATGVDTCTYHDPVPGGLDIQOLYWELSQ 430
DB |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 721 LSSPTSMOSVS-----RPGSGDASGFNT-----IFPSISQ 752
DB |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 431 LTHGVTLQPLVYLDRLSFLINGYAPONLSIRGEYQINFIIV-----NMNLSP 478
DB |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 753 SSGDET--SGYTISSNSNSONASASEPOTAFSSSSSATPTTOSISITSVSOSMNSYS 810
DB |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 479 DPTSSYITLLRDIOKV-----TTLKSGQLDHTPFPCV 514
DB |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 811 SPLSSASVSTSIISIIASSSTSIPLSISSIASGFDASGFTSITNGKAGSSFFALA 870
DB |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 515 TNLMDSVLVTVKALPFSNLDPSLVEQVFLDKTLN-----ASFHMLASTYQVLDI 564
DB |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 871 SNSBSGA-----SDVLSTIAKPTFKPSTNSGSGTSYSIPLSSSNBGTTSVSN 921
DB |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 565 HTTMESSVYQP--TSSSSTOHFYLNFTTNLFYSODKAPGCTTYQNRKNITDALNOL 622
DB |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 922 TVT---SSTLKPELTSVSSTASISASASNTLS---TEPKTFS---SSSTLSBSIS 972
DB |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 623 FRN-----SSTKSYFSDCOVSTFRSVN--RHHTGVSLCNFPLAR----- 663
DB |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 973 NTNSLTWKPCBSLSSTTSGLTSSSTIYPSSTRBSNNSASGSSAKSSSSTSLVQSN 1032
DB |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 664 -VDRVAIYERF-----LMTNRNGTQLQNFLLDRSSVLVDGYSPNNEPLTGN----- 710
DB |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 1033 PUKTVVSLSYKTKTSKISLVKPKTYVTVDVETNVVVQTHVSVYEHSTSTVHTWHT 1092
DB |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 711 ---ADIQSGGRSLGEPFEQLIS 733
DB |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 1093 YEVADVEASTKTSRPSAKSKRNAVS 1118
DB |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

```

RESULT 9
T34235
hypothetical protein F22A3.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004

C:Accession: T34235
R:Hallworth, K.
submitted to the EMBL Data Library, November 1995
A:Description: The sequence of C. elegans coamid F22A3.
A:Reference number: 221492
A:Accession: T34235
A:Species: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-532 <HAL>
A:Cross-references: UNIPROT:O19695; UNIPARC:UPI000017B91C; EMBL:U41547; PDB:AAA83194.1;
C:Gene: CESP:F22A3.1
A:Introns: 55/1; 71/1; 99/1; 148/3; 176/3; 190/1; 239/1; 303/1; 331/3; 357/1; 391/1; 511
F:446-527/Domain: etc DNA-binding domain homology <ETS>

Query Match 3.0%; Score 118; DB 2; Length 532;

Best Local Similarity 20.5%; Pred. No. 2.7; Indels 162; Gaps 25;
Matches 101; Conservative 59; Mismatches 171;

```

QY 255 DLLC-TYLQPLSG-----PGLPIKQVHELSQCTHGITRLGYPYSLDKDSLILYNGYNEP 306
DB |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 9 DFLCSTGMSGIRKKQASRQISRESVFGAKRCAVAARAP----- 52
DB |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 307 GPDEPTTPKPAITPLPLSEATTMGYHLKTLTN---FTISNLQYSPDMKG--SATF 361
DB |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 53 -PEEPFPPTPTLLFKLSERQ---FH---GLNCPAFPPNNLEKYGSGCKSDSSSM 103
DB |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 362 NSTEVLQHLRLP-----FOKSMGPFYLCQLISLRPEKGAATGVDT 406
DB |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 104 NGT-GSVGRNLSLPEPHSGTSTASTPFAVSEF-PF-----DDLLFGIDQ 148
DB |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 407 TCTYAPDPVGP--GDIQOLYWELSQLTHGVTLQPLVYLDRLSFLINGYAPONLSIRGE 463
DB |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 149 VNNVPRHMDMCMPIQPIENNR-----FSKADHSFTVKNIEINI----- 192
DB |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 464 YQINFIHVMNLSNDPTSSSEYITLLRD-----IQDKVTLYKQSQLDHTF----- 509
DB |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 193 -LNNF-NVPEIAQNGDLDTQIDYRDLRLHLIDISTCAKGLPDDFYLMSEHGA 249
DB |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 510 -----RFLCVTLN-----TMSVLYTVKALPFSNLD 535
DB |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 250 RWINEMCQFNLOPPRNCISITGIDLGMSQKDFEMILPAGGDTLHAQLQVMTAAENYHP 309
DB |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 536 PSLVE---QVFLDKTLNLSFHWLGST-YQVLDIHTEMESSVY-----OPTSSSSTOHF 585
DB |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 310 PTVQSSGTAENNMGSITNMLASTNQTNNMAAENNHFFPNGNGYPMMSMS--F 367
DB |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 586 YLNFITTLNLPYQDKAPGCTTYQNRKNIEDALNQLFRNSSIKSYFSDCOVSTFRSVN 645
DB |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 368 FQGGTVLPSPNSDTSNSGSGQ-DNMDDDDIDLMNN-----SNCGPSNFFHNOG 415
DB |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 646 RHHTGVDSLGNFS 658
DB |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 416 YNMSPIDAMNGS 428
DB |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

```

RESULT 10
A38926
DNA-binding protein cl (D) - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C:Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 09-Jul-2004
C:Accession: A38926; A38917; 512122; 512769
R:Slusarek, D.
submitted to GenBank, October 1990
A:Reference number: A38926
A:Accession: A38926
A:Molecule type: mRNA
A:Residues: 1-1377 <SLU>
A:Cross-references: UNIPROT:P19538; UNIPARC:UPI000016B841; GB:X54360; NID:97732; PID:977
R:Orenic, T.V.; Slusarek, D.C.; Kroll, K.L.; Holmgren, R.A.
Genes Dev. 4, 1053-1067, 1990
A:Title: Cloning and characterization of the segment polarity gene Cubitus interruptus 4
A:Reference number: A35817; MUID:90346286; PMID:2166702

A:Accession: A35817
A:Molecule type: mRNA
A:Residues: 'MKISFIRSIQKILTCU', 114-1151, 'K', 1152-1377 <ORE>
A:Cross-references: UNIPARC:UPI000017BE89; GB:X54360
C:Genetics:
A:Gene: FlyBase:cl
A:Cross-references: FlyBase:FBgn0004859
C:Keywords: DNA binding; zinc finger

Query Match 3.0%; Score 117.5; DB 2; Length 1377;
Best Local Similarity 19.6%; Pred. No. 14;
Matches 132; Conservative 83; Mismatches 224; Indels 233; Gaps 31;

```
QY 166 VNGFTHRSSVPTSTGVSEEPFTLNFTNN-----LRTMAMGQCGSL 209
DB 675 VNGVRASDSLLTTSPPDLAE-----NLNDDGMNCDDVDVADLPVLRMAMVINGNMS 729
QY 210 KFNITDVMKHLSPLEFORSSLGARYGCRVIALRSVNGAETRVLDLCTYLOPLSGPL 269
DB 730 ASTIGGSVLAR-----QRRFG-----RLQTKGINSSTMICNIPF----- 764
QY 270 PIKQVHEHLSQOQTHGTRIGPYSLDKDSLNGYNEBGPDEPPTPKPATFPLPULSEAT 329
DB 765 -----SNRTFGISELNQRI TELKM-----EBGTDAEIKIPKLPTTTIGSISDP 808
QY 330 TAMGYNHAKTLITLNFITSNLOQYSPDMKGSATFNSTEGVLOHLRPLFOKSMGPFIYGCQ 389
DB 809 LQ-----NOTSFNNTVSNKQ-----GIVSGSIGQOPRRDSQNSTASTYYSMQ 851
QY 390 LISLRPEKDGATGVDP-----TCTYHPDPVGGDLIQQLYWEISQLTHG----- 434
DB 852 ---SRSSQSQVSIFPMRNPSCNSTASTSY-DPISGCSRRS-----SQMSNGANCNF 903
QY 435 VTQGLGFVLDKDS-----LPINGVAPONLSIRGEYOQINFHIVMNLSPDPISSEVI-TL 488
DB 904 TSTSGLPVLANKESNKSILNACIN---KENIGVQG-----VGIYNSLSPPPSHILATN 953
QY 489 LRPIQDKVTLLYKGSQ-----504
DB 954 LKRLQRDSSEYHNFSGRFVSPYSMSHLIKNNKPVGENEFPDKAISMARQTDPPVNNIN 1013
QY 505 ---LHDPFRFCV-----TNLTMSVLTVVKALFSSNLDPSLVEQ 541
DB 1014 LDELITNISRSTTPHSDIVNGKTNINIASLKNONLAKDLFTVISIKADAMATSDQHPNER 1073
QY 542 VFPLDKT-----LNASFMLGSTYQVLDI--HYTE-MESSYQOPTSSSTQ 583
DB 1074 INLDEVABELLPDEMIQYLNLVKDDTNHLEKEHQAVPVGSNVEGTTIASNHYRQGSNT--- 1130
QY 584 HFLPLNFTITNLPYSQDKAQPGTTNYQANKENIEDALNQLFRNSSIKSYSPDCOVSTFRSV 643
DB 1131 -YTNKQILTPPSNVD-IQNTTFTVQDK-----FAMTAVGGSFQORELSL-AV 1177
QY 644 PNHH-HTGVDSLNCFSPLARVD-----RVAIYEEFLRMTRNGTQLONFLDLS-----S 692
DB 1178 PNHHGAKKCSFHHSQSKYNNMTDIGSKQSAALPSAHORQTEKSN--YNOIIDSMTSLPE 1235
QY 693 VLVDGYSPNENE 704
DB 1236 LWDSTIYP-RNE 1246
```

RESULT 11
A44140
cellulose-binding protein A precursor - Clostridium cellulovorans
C:Species: Clostridium cellulovorans
C>Date: 27-Jun-1994 #sequence_revision 27-Jun-1994 #text_change 09-Jul-2004
C:Accession: A44140
R/Shoheyov, O.; Takagi, M.; Goldstein, M.A.; Doi, R.H.
Proc. Natl. Acad. Sci. U.S.A. 89, 3483-3487, 1992
A>Title: Primary sequence analysis of Clostridium cellulovorans cellulose binding protei
A:Reference number: A44140; MUID:92228810; PMID:1565642
A:Accession: A44140

A>Status: Preliminary
A:Molecule type: genomic RNA
A:Residues: 1-1848 <SHO>
A:Cross-references: UNIPROT:P38058; UNIPARC:UPI00001271G5; GB:M73817; NID:G144748; PIDN
C:Genetics:
A:Gene: cbpa

Query Match 2.9%; Score 116; DB 2; Length 1848;
Best Local Similarity 17.7%; Pred. No. 26;
Matches 133; Conservative 92; Mismatches 244; Indels 284; Gaps 30;

```
QY 14 FTHRSVSTSTSPGTPV-----YIGASK-----TPASIFGPSA-----S 49
DB 154 YQTNVYSDDASSSTVVPKVTGIGAKVIGTARPPVPSIINPTSATPDKNVTYKQA 213
QY 50 HLLILFTLN---FTTNLRYEENMPSGRKNTTERVLQ---LRLPLFKNTSVGPLY 101
DB 214 DVKTWTTLNGNTEFKITTDAN-----GTALNASTVSVSGNDVTISKAYIAKQSVGTTT 266
QY 102 SGCRLTLNLPK-----DGEATGVDAICTHRPDPGPGLDRE 138
DB 267 LNFNFSAGNPQKLVITVVDTPVEAVTATIGKQVNAGETVAVFVNLIKVP---AAGLATI 323
QY 139 QLYELSQLTHSITE-----LGPYTLDRDLY-- 165
DB 324 ELPLTDSASLEVVSTIADIVLNPSNFSSTVSSGTTKLFPLDPLGSQLIKDGVFAT 383
QY 166 -----VNGFTHR-SSVPTSTGVV---SEEPFTLNFTNNLRVMADMGQPSGLKEN 212
DB 384 IFPKAKAIGTAKVTSVLAGTPVVGDAQLDEKPCAVN-----PGVTIN 429
QY 213 ITDVMKHLSPLEFORSSLGARYTGCRVIALSVKN---GAEITVDLCTYLOPLS--- 265
DB 430 PIDNRQGISVG---TATYKAGEIAAVPVLTVSPSTGATGAEOVSFPATLLEVASVTA 485
QY 266 -----GP-----GLPIKQVHEHLSQOQTHGTRIGPYSLDKDSLNGYNEBGPDEP 311
DB 486 GDIVLNPTVNFSTVGNVYIKLFLD-----DPLGSQLSKDGVPT-IPFKAKAVT 536
QY 312 PTPPKPATFPLPULSEATTAMGYNHAKTLITLNFITSNLOQYSPDMKGSATFNSTEGVLOHL 371
DB 537 SFVTTFTVTVSGTFVFDGLTAEVQSKT-----AAGSVTINIDPILPEPT 580
QY 372 LKPL---FOKSSMGPFYLGCOLISLRPEKDGATGVDTCTYHPDPVGGDLIQQLYWEI 428
DB 581 ISPVATTFPKK-----APADATMT----- 601
QY 429 SQLTHGVTQGLFVLDKSLFINGVAPONLS--IRGEYOQINFHIV----- 471
DB 602 -----LNGYTFNGITGLTSDYISIGNVVKISQAYILAKQVGYDL 640
QY 472 --NMNLSPDPT-SSEYITLBDIDQKVT-----TLYKG-----SQL 505
DB 641 TLTFNFSNNKATATLAVSINDAPRTVATGTATVNAGETVAVPVTLSSNVGISSTABL 700
QY 506 HDTEFPCV---VTNLMDSVLTVVKALFSSNLDPSLVEQVFLDKTLNASTHMLGSTYQVLD 563
DB 701 QLSFDATLLEVVISITAGDIVLNPSVNFSSVNVGSIKLFLLDPLDLSQSLSKDGVPATIN 760
QY 564 IHTHESSVYQPTSSSTQHPLNFTINLPV 596
DB 761 FPAKSVSTVTFVVKVSGTFVP-ADGTLAELSY 792
```

RESULT 12
F69009
probable membrane protein MTH1074 - Methanobacterium thermoautotrophicum (strain Delta)
C:Species: Methanobacterium thermoautotrophicum
C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C:Accession: F69009
R/Smith, D.R.; Douchette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
; Oiu, D.; Spadeford, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N
ki, S.; Church, G.M.; Daniele, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.

J. Bacteriol. 179, 7135-7155, 1997
A:Title: Complete genome sequence of *Methanobacterium thermoautotrophicum* Delta H: funct
A:Reference number: A69000; MUID:98037514; PMID:9371463
A:Accession: F69009
A:Staccus: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1474 <MTH>
A:Cross-references: UNIPROT:Q27146; UNIPARC:UPI00000668F; GB:AE000878; GB:AE000666; NID
A:Experimental source: strain Delta H
C:Genetics:
A:Gene: MTH074
C:Keywords: duplication

Query Match 2.9%; Score 115.5; DB 2; Length 1474;
Best Local Similarity 20.5%; Pred. No. 21;
Matches 158; Conservative 92; Mismatches 306; Indels 215; Gaps 38;

QY 26 PGTPT-VYLGASTPASPISGPAASHLLFTLN-FTTNLYEENMMGSKRFTTERV 83
DB 393 PNNATGVYVTDLLPQLSFSVASASRGVNSTGWTIGNLEFETV-----TLNITATV 447
QY 84 LQGLRLPLKMTSVGLVSGCRIT--LAREKDEATGVDAICTHRPDTGRLDREQLY 141
DB 448 -----TATGATVNNANVTGDFEDDM--ANNYASALNPPASDLTIDK--- 489
QY 142 LELSLQTHSITELGPTLDRDSLTVNGFTHRSEVPTTSGVSEEPF----- 188
DB 490 -----SVNNRBP-VGENIOTTTVSNRG--PNNAGVVEVDLPGALPISATPK 538
QY 189 -----TLNF-TINNLRVADMGOPGSLK-FNITDVMKHLPLP--QRSS 230
DB 539 GSYGWTGNNVGLTNVLEIATLTIIARVNATGSLTNFANIT-----SPNFDPNDDNN 590
QY 231 LGARTYGCRIALRSYKNAETVVD-----LCTYIQPLSGPLPIKQVPHF-LSQ 281
DB 591 DTAENVGIPVADLLIKQVSDPRPDYGVVTFVAVTNLGPSTNATVATDILSPGLVYL 650
QY 282 TGTGRTLPYSLDKOSLYNGVNEPQDEPPTPKPATTFPLPISATTAAMGHLKTLTL 341
DB 651 SHVTVQ-GTYNATGWTGYGALNY-----NAGSALMLVTVNITGDSNNTV-- 695
QY 342 NPTISNLQYSPDMKGSA--TFNSTEGLVQHLRLPLFKSSMGPFYLGQQLSLAREKDG 399
DB 696 --SVTGENRDPDRNTNDVASTLNAVSADLS-----IQKTVD-----REVINN 735
QY 400 AATGVDTTCTYHPDPVPGGLDT--QQLYBELSQLTHGVTQQLGY---VLDRDSLFINQ 453
DB 736 GRTATPTVIVRNAGPPTPSNVVSDLLPAGLSIISYTVQGSFNTTGVMEVGLPALPQ 795
QY 454 APQNLISIRGE--YOINPHIVMNLNPNPTSEYITLL-----RDIQDKVTLLYKGS 503
DB 796 ATTLVLRATQAGFQTN--LVNVSSELPPDLPGDNDAVTVDRPADVKITTVNTAP 853
QY 504 QLHDTFRFCL-VTNLTMDSVLYTVKALFSSNLDPSSLVEQVFLDKTLNASHFWLGSYQYL 562
DB 854 DFDYTVVFYITVNLDPDATV-----RT----- 878
QY 563 DIHVTMESSVVOPTSSSTQHFYLNFTTNLFYSQDKAPGTTNQRKNIEDLNQL 622
DB 879 ---VDTMPGQLVYQSHSASAGIYFEPNV---WTYDLSLAPGASE-TLNTTVLVNATGCM 930
QY 623 FRNSISKYSFSDCOVTFPSVRHHTGVDSLGNFSLARVDNRVAIYEEFLMTANGNQ 682
DB 931 INTVASTSTTEYDDLT-----NNHAG--LNN---AEAVADIAQVKTLTTPPNNQ 977
QY 683 LQNFLLDRSSVLVDGYSPNREPLTQNSAD-----IQHSGGRSSLEG 724
DB 978 ITNFT-----VYTNNGPN---DATGAATVDILPGLGLSHSASQGTFA 1020

RESULT 13
A54280
cell differentiation protein lagc precursor - slime mold (Dictyostelium discoideum)

N:Alternate names: loose aggregate C
C:Species: Dictyostelium discoideum
C:Date: 09-Sep-1994 #sequence_revision 09-Sep-1994 #text_change 09-Jul-2004
C:Accession: A54280
R:Idynes, J.L.; Clark, A.M.; Shaulsky, G.; Kupea, A.; Loomis, W.F.; Firtel, R.A.
Genes Dev. 8, 948-958, 1994
A:Title: lagc is required for cell-cell interactions that are essential for cell-type di
A:Reference number: A54280; MUID:95011593; PMID:7926779
A:Accession: A54280
A:Staccus: preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-888 <DYN>
A:Cross-references: UNIPROT:P42523; UNIPARC:UPI000012E1FF; GB:U09478; NID:9495172; PIDN:
A:Note: authors translated the codon TCA for residue 34 as Ieu
C:Genetics:
A:Gene: lagc
C:Superfamily: Dictyostelium discoideum cell differentiation protein lagc
C:Keywords: transmembrane protein

Query Match 2.9%; Score 115; DB 2; Length 888;
Best Local Similarity 19.6%; Pred. No. 10;
Matches 157; Conservative 109; Mismatches 235; Indels 298; Gaps 47;

QY 110 RPEKGE--ATG-----VDAICTHRP-----DPT-----GPGLD 136
DB 130 KPKTSGEDIVTGSYLRLFGGEPFLINSIDVKNPVGKGFSDPFCNNITVTFPPGSG 189
QY 137 RQQLYBELSQLHSTELG---PYTLD-----RSLVNG--PTHRSSVPT 177
DB 190 KRLLYD-----ETGDNVPFVSYESPIISSVSDBSKQIITINGDNFTDQDLVY 240
QY 178 TSTGVASEEPFLNFTI-----NNLRVADMGOPGSLKFNIT-- 214
DB 241 SFDGI--DQP--NFIISNMHQIQVNNYNR--DEQPMVNTVDEVSIEKQYHCF 291
QY 215 -----DNVMKHL-----LSPLFKRSSIGARTGCRVIALASVKNQ 249
DB 292 PAITISISVSNHLGIVITIKGEKLSLTNLYLTFSI--TIGDYK-----VLIKSTTE 343
QY 250 AETRYDLCITYQPLSGPLPIKQVPH-----ELSQQTHGI----- 285
DB 344 LKCKLD--ANELGKXKLPVNVNFGGCDSTSPNGVSTFYNIPLTSSRGSYSGIVTLIG 398
QY 286 TELGPPSYLKDLSLYNG-----YNERGPDEPPTPKPATTFPLPISLEAT-----TMM 332
DB 399 TNLGTMNESSIDLYGEGIKNTNISQFNVSSE-----KSYFELPHLRCSFNINFRS 453
QY 333 GYHLKTLTLNFTIS-NLQYSPDMKGSAATFNSBGLVQHLRLPLFKSSMGPFYLGQQL 391
DB 454 GITAKTLTISASLSVNVINRPTV-----SNGIL-----NIEIYMDCTIS 493
QY 392 SLRPE-----KQGAATGVDTTCTYHPDPVPGGLDIQQLYBELSQLTHGVTQ 437
DB 494 SSAPSLTVGDSSASAPCSIPSSNSSYETTC--PTPYGTGIN-KQIFKLNSET----- 544
QY 438 LGFVYVLDKSLFLINGAP--ONLSIRGE-YOINPHIVMNLN--NPDPTSEY 485
DB 545 -----VSDQFS-----YAPPEVENRTISDQTNIELHGNFGASTLKYKLVNGSDISSF 594
QY 486 ITLARDIOD-----KVTTLYK-----GSQLHDTFRFCL--VTNLTMDSV-- 522
DB 595 ---IQLELDHQTLIKLDSYENGPNITVDGVMYMSLFYTLPLVLYRITNDKNKTLAC 651
QY 523 -LTVYKALFSSNLDSVLEQVFLDKTLNASHFWLGSYQYLVDVNIHTEMSSVVOPT----- 577
DB 652 GIIYVS--GKQLTSLDKFKNVNVSNKNTTVFAQDEKILLVRDESRSSSLFVTFICV 708
QY 578 -GSSSTQHFYLNFTTNLFYSQDKAPGTTNQRKNIEDLNQLFRSSISKYSF-- 633
DB 709 RSGPSTTLTYIKPMISEITIKIEKGLAIIRG-YSTFDILNASLTVSSSEVPISCHL 767
QY 634 DCQVS-----TFSSVPRHHTGVDSL-CNPSPLARVDNRVAIYEEFLMRNQTQLQ- 684

Db 768 ECSLSPNEILDSDSETNTNTDCLSCHSGSSVAKNTSGV-----LYLFFNSTSFQY 821
QY 685 NPTLDRSSVAVDGYSPRN 703
Db 822 NVTIEIKT-----SPSPN 835

RESULT 14

A35364
carcinoembryonic antigen-related protein (clone rncGM1) - rat
C/Species: Rattus norvegicus (Norway rat)
C/Date: 17-Aug-1990 #sequence_revision 17-Aug-1990 #text_change 09-Jul-2004
C/Accession: A35364
R/Rebstock: S.; Lucas, K.; Thompson, J.A.; Zimmermann, W.
J. Biol. Chem. 265, 7872-7879, 1990
A/Title: CDNA and gene analyses imply a novel structure for a rat carcinoembryonic antigen
A/Reference number: A35364; MUID:90243655; PMID:2335509
A/Accession: A35364
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-709 <REB>
A/Cross-references: UNIPROT:Q63111; UNIPARC:UPI0000E81AF; GB:M32474; NID:G203430; PIDN:
C/Superfamily: carcinoembryonic antigen-related protein rncGM1; carcinoembryonic antigen
F:1-137/Domain: carcinoembryonic antigen precursor amino-terminal homology <CEA1>
F:168-217/Domain: immunoglobulin homology <IMM1>
F:242-378/Domain: carcinoembryonic antigen precursor amino-terminal homology <CEA2>
F:480-612/Domain: carcinoembryonic antigen precursor amino-terminal homology <CEA3>
F:633-690/Domain: immunoglobulin homology <IMM2>

Query Match 2.9%; Score 114.5; DB 2; Length 709;
Best Local Similarity 20.0%; Pred. No. 7.6;
Matches 136; Conservative 80; Mismatches 263; Indels 201; Gaps 30;

QY 134 GLDREQLYLELSQTHSITEIGPYTLDRSLVYNGFTHRSSVPTTSGVSEBPTLNT 193
Db 70 GLTNMSGVALYSLTYVTVTGPHSGRETLVSGSLMIGNVQKDTGYT----- 120
QY 194 INNLRYADMGQGSGLKFNITDVMKHLSPFORSSLGARYTCRYALRSVNGAETR 253
Db 121 --LRTISNMG-----IVSNTSLH-LHVFSTLTCGAAATSAQ--LSIESVPTSISSG 168
QY 254 VDLICTYLOPLSGRPL--IKOVFH-----ELSQGTHGITRL--GYSLKDQSLY 299
Db 169 ESAL-----LAAHNPENIRAFWYKGAIVFKDLVARYVIGTNSSVPEPAHSGRTMY 222
QY 300 LMG-----YNEBGPDEPTTPKPAITFLPLSEATTAMGY-HLKTLLNFTISMLQYSPDM 354
Db 223 SNCSLLQNTVRNDAGYTLKTLST-----DLKTEIAVYQLQVDTCFMSYAGPPTSAGL 276
QY 355 GKGSATFNSTEG---VLQHLRP-----LFQKSMGPFYLGCOLISLRPEKD 398
Db 277 TVESAPTSVAEGASVLLVNLNLPENIRAFWYKGVILFKDLVARYVIGTNSSVGLGPAHS 336
QY 399 GAATGVDTTCTHPDPVGPGLDLOQLYWEISQLTHGTC--LGPVYLDROSLINGAPQ 456
Db 337 GRET-----MYSNGSLQLQNTVRNDAGYTLRTSLDLSKAV-- 373
QY 457 NLSIRGEYQINFIIVNNILSNPDTSEYITLLRDIDQKVT--LYKSQLHDTFRFLVY 515
Db 374 ---VHQQLQVN-----TSSCCDPLTPALLTI-----DVPRIHAAG----- 406
QY 516 NLTMDSVLTVALFSSNLDPSLVEQVLDKTLNASFHVLGTYQLVDIVHTEMES-- 572
Db 407 ---ESVLLQVR-----NLPEDL-----RMFIWFKSVYTSQIFKIAYSRAINY 446
QY 573 VYQPTSSSTQHYLNFTINLPYSODKAPGTTNVRNKRNEEDLNOLFSSSIKSYF 632
Db 447 VFGPAPASGREYTYTNGSL-----LLODATERKDTGLY-----TLQITIRNFKLETAAH 493
QY 633 SDCOVST-----FRSVNRRHHTGVDSLGNFSLARVDRVAIVEEFLRMTENG 681
Db 494 VQVSVHTCVAPSTTGQVLIVTSVPPNVVEGQDVL-----LLVHNMP 534

QY 682 QLQNFPLDRSSVAVDGYSPRNPEPLTNGSADI--CHSGRSSLEGRPFEOKLISEEDLNM 739
Db 535 NLSQFSWYKGVAILVNHHEISRN--IISNRSTLGPAPASGREYITNSGSLILHNATEEDNG 593
QY 740 HT-----GHHHHH 748
Db 594 YTLMTVNRHSETOGIVHH 613

RESULT 15

T07015
Cf-4A protein - tomato
C/Species: Lycopersicon esculentum (tomato)
C/Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 09-Jul-2004
C/Accession: T07015
R/Takken, F.L.; Schipper, D.; Nijkamp, H.J.; Hill, J.
Plant J. 14, 401-411, 1998
A/Title: Identification and De-tagged isolation of a new gene at the Cf-4 locus of toma
A/Reference number: Z15863; MUID:98335213; PMID:9670557
A/Accession: T07015
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-855 <TAK>
A/Cross-references: UNIPROT:O50024; UNIPARC:UPI00009FC38; EMBL:Y12640; NID:e1289424; P
C/Genetics:
A/Experimental source: strain Cf-4; isolate MM-Cf-4
A/Map position: 1

Query Match 2.9%; Score 113.5; DB 2; Length 855;
Best Local Similarity 19.6%; Pred. No. 12;
Matches 149; Conservative 111; Mismatches 209; Indels 293; Gaps 43;

QY 54 LFTLNTTINLRYEN-----MWPGRKNTTERRVLO-----GLRLPRKMTS 96
Db 42 MFTINPMAINYCYDRRTLSMNKSTSCSMDGVHCDETQVLELDGCSQLQKPFHSNS 101
QY 97 VGPVYSGCLTLTPKPKDEATGVDAICTHRDPPTGPGDREQLYLELSQTHSITEIGP 156
Db 102 LQQLSLMKLIDL--SSNDFTG-----SPISP-----KFGFSDLTN----- 135
QY 157 YTLTD-RDSLTVNGFTHRSSVPTTSGVSEB-----PFTLNTFI 194
Db 136 --LDLSDSNF-----TGVIPELSHLSKHLAVLRISDYKSLGHNELLL 179
QY 195 NNILRYADMG-QGSLKFNITDVMKHLSPFORSSLGARYTCRYALRSVNGAETR 253
Db 180 KNLQTLRELHLSVNISSITPSNFSFHL-----TNLRSLSTELR----- 218
QY 254 VDLICTYLOPLSGRPLPIKOVFH---ELSQGTHG---YRLGPHYSLDKS-----LYLNG 302
Db 219 -----GVLPRVFHLSNLELDLISYNQLTVRFPTTITWNSASLVLKYLIS 264
QY 303 YNEBGP--PDEPTTPKPAITFLPLSEA-----TTAMGYHLKTLTLLFTISNLQYSPDMGK 356
Db 265 VNIAGNIP-----SPSYLTALHELDVYTNISGPIPKPL--WNLTNLE-SLDL-- 310
QY 357 GSATFNSTEGVLOHLRPLFOKSMGPFYLGCOLISLRPEKGAATGVDTCTYHHPDV 416
Db 311 ---DVNHLEGPILQOL--PIFEK-----LNSL-----TLGNND 339
QY 417 PGDLIDQLYWEISQLTHGVTQLGFFVYLDROSLFINGYAPONIS---INGEY---QINF 468
Db 340 GGLLEFLSPNRSWTOLEE-----LDPSNSNLTCPIPNSVSGLRNLSLYLSSNNLNG 390
QY 469 HLYNM-----NLSNPDTSEYITLLRDIDQKVTTL--YKGSOL-----HDTFRP 511
Db 391 STPSWLFPLDPLSLSLDLSNNTPSGKIQEFKSTLISVTLKQNDLKKPINSLLNQSSLOP 450
QY 512 CLVT--NLV--MDSVLTVALF-----SSNLDPSLVEQV-----LDRT 547
Db 451 LLLSHNINISGHSSSICNLKILMVLIDGNNIEGTIPQCVVERNEYLSHLDSNNRLSGT 510

